SEQ UPDATE 1-3, 55-56
(28-29 - Not Reported)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 36.7714 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	26	21	AAB37001	Bcl2 polypeptide B
2	138	100.0	26	21	AAB37002	Bcl2 polypeptide B
3	138	100.0	27	21	AAB37003	Bcl2 polypeptide B
4	138	100.0	27	21	AAB37056	Bcl2 polypeptide B
5	138	100.0	28	21	AAB37055	Bcl2 polypeptide B
6	138	100.0	162	22	AAB70370	Shorter murine BAD
7	138	100.0	204	17	AAR95168	bcl-x(L)/bcl-2 ass
8	138	100.0	204	19	AAW61315	Murine BCL-XL/BCL-
9	138	100.0	204	19	AAW61316	Mutant BCL-XL/BCL-
10	138	100.0	204	19	AAW61317	Mutant BCL-XL/BCL-
11	138	100.0	204	19	AAW61318	Mutant BCL-XL/BCL-
12	138	100.0	204	19	AAW58832	Murine BAD protein
13	138	100.0	204	22	AAB70369	Longer murine BAD
14	138	100.0	204	24	ABR39082	Murine BAD protein
15	138	100.0	567	22	AAU00220	Bad-DTTR apoptosis
16	114	82.6	24	23	AAU78627	Human Bad peptide
17	114	82.6	25	23	ABP56161	PTPC-interacting T
18	114	82.6	25	23	ABG78484	Mutant Bcl2 compet
19	114	82.6	25	23	ABG78493	Mutant Bcl2 compet
20	114	82.6	25	23	AAU78610	Human Bad peptide
21	114	82.6	25	23	AAU78620	Human Bad peptide
22	114	82.6	166	18	AAW32476	BBC6 protein for r
23	114	82.6	168	19	AAW55779	Human Bcl-xL/Bcl-2
24	114	82.6	168	21	AAB13512	Human cell prolife
25	114	82.6	168	22	AAB70368	Human BAD mutant a
26	114	82.6	168	22	AAB48287	Human Bad protein.
27	114	82.6	168	22	AAG67688	Amino acid sequenc
28	114	82.6	168	24	ABR39081	Human BAD protein
29	114	82.6	201	23	ABP41630	Human ovarian anti
30	113	81.9	23	17	AAR95166	bcl-x(L)/bcl-2 ass
31	111	80.4	25	23	ABG78490	Mutant Bcl2 compet
32	111	80.4	25	23	AAU78617	Human Bad peptide
33	110	79.7	25	23	ABG78488	Mutant Bcl2 compet
34	110	79.7	25	23	ABG78489	Mutant Bcl2 compet
35	110	79.7	25	23	AAU78615	Human Bad peptide
36	110	79.7	25	23	AAU78616	Human Bad peptide
37	109	79.0	23	23	AAU78628	Human Bad peptide
38	109	79.0	25	23	ABG78486	Mutant Bcl2 compet
39	109	79.0	25	23	ABG78492	Mutant Bcl2 compet
40	109	79.0	25	23	ABG78497	Mutant Bcl2 compet
41	109	79.0	25	23	AAU78612	Human Bad peptide
42	109	79.0	25	23	AAU78619	Human Bad peptide
43	109	79.0	25	23	AAU78624	Human Bad peptide
44	108	78.3	25	23	ABG78485	Mutant Bcl2 compet
45	108	78.3	25	23	AAU78611	Human Bad peptide

ALIGNMENTS

RESULT 1
AAB37001
ID AAB37001 standard; peptide; 26 AA.
XX
AC AAB37001;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #1.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.

XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 2
AAB37002
ID AAB37002 standard; peptide; 26 AA.
XX
AC AAB37002;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #2.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 2e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 3
AAB37003
ID AAB37003 standard; peptide; 27 AA.
XX
AC AAB37003;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #3.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.

XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 138; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 4
AAB37056
ID AAB37056 standard; peptide; 27 AA.
XX
AC AAB37056;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #56.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 138; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 5
AAB37055
ID AAB37055 standard; peptide; 28 AA.
XX
AC AAB37055;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #55.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.

XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 28 AA;

Query Match 100.0%; Score 138; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 6
AAB70370
ID AAB70370 standard; protein; 162 AA.
XX
AC AAB70370;
XX
DT 02-MAY-2001 (first entry)
XX
DE Shorter murine BAD mutant amino acid sequence SEQ ID NO:3.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 7; Page 148-149; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed shorter murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 138; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 1.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 98 NLWAAQRYGRELRRMSDEFEGSFKGL 123

RESULT 7
AAR95168
ID AAR95168 standard; Protein; 204 AA.
XX
AC AAR95168;
XX
DT 06-JAN-1997 (first entry)
XX
DE bcl-x(L)/bcl-2 associated death promoter protein.
XX
KW Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS Mus musculus.
XX

```
FH   Key             Location/Qualifiers
FT   Region          147..149
FT                   /note= "BH1 conserved amino acids"
FT   Region          191..192
FT                   /note= "BH2 conserved amino acids"
FT   Domain          38..61
FT                   /note= "PEST sequence"
FT   Domain          111..130
FT                   /note= "PEST sequence"
XX
PN   WO9613614-A1.
XX
PD   09-MAY-1996.
XX
PF   31-OCT-1995;    95WO-US14246.
XX
PR   31-OCT-1994;    94US-0333565.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1996-251465/25.
DR   N-PSDB; AAT29479.
XX
PT   Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT   useful to treat neoplasia and apoptosis and to identify agents
PT   inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS   Claim 3; Fig 1; 130pp; English.
XX
CC   This sequence represents the murine bcl-x(L)/bcl-2 associated death
CC   promoter (Bad) gene.  Bad is a 22.1 kD protein which interacts with
CC   bcl-2 and bcl-x proteins and regulates cell death.  It has homology
CC   to the bcl-2-related family clustered in the BH1 and BH2 domain.  Bad
CC   has been found to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid
CC   assays and in vivo in mammalian cells.  Overexpressed Bad counters the
CC   death inhibitory activity of bcl-x(L), but is much less effective at
CC   countering the death inhibitory activity of bcl-2.  Bad expression can
CC   accelerate apoptotic cell death induced by cytokine deprivation in an
CC   IL-3 dependent cell line expressing bcl-x(L), and its also counters the
CC   death repressor activity of bcl-x(L).  Bad competes with Bax for binding
CC   to bcl-x(L).  Bad may be used to identify agents which inhibit its
CC   binding to bcl-2 or bcl-x(L) to form heterodimers.  Such agents may be
CC   used to treat neurodegenerative diseases, immunodeficiency diseases,
CC   e.g. AIDS, senescence or ischaemia.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 17;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 8
AAW61315
ID   AAW61315 standard; Protein; 204 AA.
XX
AC   AAW61315;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Murine BCL-XL/BCL-2 associated cell death regulator.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;    97WO-US19175.
XX
PR   18-OCT-1996;    96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27833.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 1; Fig 10; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence is the murine BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 9
AAW61316
ID   AAW61316 standard; Protein; 204 AA.
XX
AC   AAW61316;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #1.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
```

```
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27834.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 59; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 10
AAW61317
ID   AAW61317 standard; Protein; 204 AA.
XX
AC   AAW61317;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #2.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27835.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 60; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 11
AAW61318
ID   AAW61318 standard; Protein; 204 AA.
XX
AC   AAW61318;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #3.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
```

```
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27836.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 60-61; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 12
AAW58832
ID   AAW58832 standard; protein; 204 AA.
XX
AC   AAW58832;
XX
DT   23-JUL-1998  (first entry)
XX
DE   Murine BAD protein.
XX
KW   BAD protein; Bcl-XL/Bcl-2 associated cell death regulator; 14-3-3;
KW   serine phosphorylation; post-translational modification; apoptosis;
KW   signal transduction regulator; phosphoserine phosphatase; senescence;
KW   immunodeficiency disease; neurodegenerative disease; infertility;
KW   cancer, viral infection; lymphoproliferative condition; arthritis;
KW   inflammation; autoimmune diseases.
XX
OS   Mus sp.
XX
PN   WO9809643-A1.
XX
PD   12-MAR-1998.
XX
PF   09-SEP-1997;   97WO-US15871.
XX
PR   09-SEP-1996;   96US-0707868.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-207049/18.
XX
PT   Serine-phosphorylated Bcl-X-l/Bcl-2 Associated cell Death regulator
PT   polypeptide - useful for modulation of apoptosis associated with,
PT   e.g. cancer and immunodeficiency diseases
XX
PS   Claim 3; Fig 8; 61pp; English.
XX
CC   This sequence represents a novel serine-phosphorylated protein, BAD
CC   (Bcl-Xl/Bcl-2 associated cell death regulator). The serine residue is
CC   phosporylated in a post-translational modification and allows binding
CC   to the 14-3-3 protein which is a signal transduction regulator.
CC   Modulators of phosphorylated BAD, which act through inhibition/activation
CC   of a phosphoserine phosphatase, are useful for preventing/treating
CC   increased/decreased apoptosis in a cell. The increased apoptosis may
CC   result from immunodeficiency diseases, senescence, neurodegenerative
CC   disease, ischaemic cell death, reperfusion cell death, infertility and
CC   wound-healing.  Decreased apoptosis may result from cancer, viral
CC   infection, lymphoproliferative conditions, arthritis, infertility,
CC   inflammation and autoimmune diseases. Measuring the amount of
CC   phosphorylated compared to unphosphorylated BAD polypeptide and/or total
CC   BAD in a cell is useful for determining the apoptotic state of a cell.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 13
AAB70369
ID   AAB70369 standard; protein; 204 AA.
XX
AC   AAB70369;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Longer murine BAD mutant amino acid sequence SEQ ID NO:2.
XX
KW   Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW   immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW   cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW   immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW   immunodeficiency disease; neurodegenerative disease; viral infection;
KW   ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW   lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS   Mus musculus.
OS   Synthetic.
XX
PN   WO200110888-A1.
XX
PD   15-FEB-2001.
XX
PF   30-MAY-2000; 2000WO-US11864.
XX
PR   28-MAY-1999;   99US-0136783.
```

XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 4; Page 148; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed longer murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 204 AA;

```
  Query Match           100.0%;  Score 138;  DB 22;  Length 204;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165
```

RESULT 14
ABR39082
ID ABR39082 standard; Protein; 204 AA.
XX
AC ABR39082;
XX
DT 10-MAY-2003 (first entry)
XX
DE Murine BAD protein SEQ ID NO:4.
XX
KW Mûrine; BAD; herpes simplex virus; HSV; US3; herpes virus; apoptosis;
KW virucide; infection.
XX
OS Mus musculus.
XX
PN WO2003012049-A2.
XX
PD 13-FEB-2003.
XX
PF 31-JUL-2002; 2002WO-US24177.
XX
PR 31-JUL-2001; 2001US-308929P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Munger J, Roizman B;
XX
DR WPI; 2003-248168/24.
DR N-PSDB; ABZ81201.
XX
PT Inducing apoptosis in a cell infected with herpes simplex virus, HSV,
PT by administering to the cell, a composition comprising an agent that
PT inhibits phosphorylation of pro-apoptotic polypeptide BAD by HSV US3 -
XX
PS Claim 15; Page 168; 192pp; English.
XX
CC The present invention describes a method (M1) for inducing apoptosis in
CC a cell infected with herpes simplex virus (HSV), which comprises
CC administering to the cell, a composition having an agent that inhibits
CC phosphorylation of pro-apoptotic polypeptide BAD by HSV US3. Also
CC described is a method (M2) for treating a patient infected with HSV, by
CC administering to the patient, a composition comprising a peptide
CC comprising a sequence of 4-100 continuous amino acids of a 168 residue
CC amino acid sequence (see ABR39081), where the peptide comprises ser112,
CC ser135, or ser155, or their combinations. BAD has virucide activity.
CC M1 is useful for inducing apoptosis in a cell infected with HSV, where
CC the cell is in a human. M2 is useful for treating a patient infected
CC with HSV. The present sequence represents murine BAD, which is used in
CC the exemplification of the present invention.
XX
SQ Sequence 204 AA;

```
  Query Match           100.0%;  Score 138;  DB 24;  Length 204;
  Best Local Similarity  100.0%;  Pred. No. 1.9e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165
```

RESULT 15
AAU00220
ID AAU00220 standard; Protein; 567 AA.
XX
AC AAU00220;
XX
DT 31-MAY-2001 (first entry)
XX
DE Bad-DTTR apoptosis-modifying fusion protein.
XX
KW Mouse; Bad-DTTR; apoptosis; cancer; spinal muscular atrophy;
KW diphtheria toxin receptor binding domain; DTR; neoplasm; tumour;
KW hyper-proliferation; Alzheimer's disease; neurodegenerative disorder;
KW transient ischaemic neuronal injury; stroke; spinal cord injury;
KW Huntington's disease.
XX
OS Chimeric - Mus sp.
OS Chimeric - Corynebacterium diptheriae.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Region 3..12
FT /note= "10x histidine tag"
XX
PN WO200112661-A2.
XX
PD 22-FEB-2001.
XX
PF 15-AUG-2000; 2000WO-US22293.
XX
PR 16-AUG-1999; 99US-0149220.
XX
PA (HARD) HARVARD COLLEGE.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Youle RJ, Liu X, Collier RJ;
XX
DR WPI; 2001-218343/22.
DR N-PSDB; AAS00248.
XX

PT Novel fusion protein for modifying apoptosis in target cell and
PT reducing apoptosis after transient ischaemic neuronal injury, has two
PT domains which targets protein to a cell and modifies apoptotic response
PT of cell -
XX
PS Claim 4; Page 59-61; 65pp; English.
XX
CC The sequence represents the amino acid sequence of Bad-DTTR apoptosis-
CC modifying fusion protein comprising Bad gene sequence fused via a short
CC linker to diphtheria toxin translocation domain (DTTR). The
CC functional apoptosis-modifying fusion protein is capable of binding a
CC target cell and integrating into or crossing a cellular membrane of the
CC target cell. The apoptosis-modifying fusion protein comprises at least
CC two domains: the DTR domain, which targets the fusion protein to the
CC target cell and the Bcl-XL domain, which modifies an apoptotic response
CC of the target cell. The fusion protein is useful for modifying
CC (inhibiting or enhancing) apoptosis in a target cell, such as neuron,
CC lymphocyte, cancer, neoplasm, macrophage, epithelial, stem, tumour or
CC hyper-proliferative cell or an adipocyte. It is also useful for reducing
CC apoptosis in a subject after transient ischaemic neuronal injury,
CC especially spinal cord injury. The fusion protein may be used to treat
CC various diseases and injury conditions through inhibition or enhancement
CC of apoptotic cellular response, including neurodegenerative disorders
CC such as Alzheimer's disease, Huntington's disease, spinal muscular
CC atrophy, stroke episodes and unregulated cell growth as in tumours and
CC various cancers. The apoptosis-modifying fusion protein can be delivered
CC effectively throughout the body and targeted to selective tissue and
CC cells.
XX
SQ Sequence 567 AA;

Query Match 100.0%; Score 138; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 6.1e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        161 NLWAAQRYGRELRRMSDEFEGSFKGL 186
```

Search completed: September 15, 2003, 17:22:13
Job time : 37.7714 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 13.5571 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	204	1	US-08-333-565-2	Sequence 2, Appli
2	138	100.0	204	2	US-08-661-479-2	Sequence 2, Appli
3	138	100.0	204	2	US-08-733-505A-1	Sequence 1, Appli
4	138	100.0	204	2	US-08-733-505A-12	Sequence 12, Appl
5	138	100.0	204	2	US-08-733-505A-13	Sequence 13, Appl
6	138	100.0	204	2	US-08-733-505A-14	Sequence 14, Appl
7	135	97.8	204	2	US-08-717-123-3	Sequence 3, Appli
8	135	97.8	204	4	US-09-375-257-3	Sequence 3, Appli
9	114	82.6	166	1	US-08-665-617-2	Sequence 2, Appli
10	114	82.6	168	2	US-08-717-123-2	Sequence 2, Appli
11	114	82.6	168	3	US-08-985-335-1	Sequence 1, Appli
12	114	82.6	168	3	US-08-985-335-7	Sequence 7, Appli
13	114	82.6	168	3	US-09-410-372-1	Sequence 1, Appli
14	114	82.6	168	3	US-09-410-372-7	Sequence 7, Appli
15	114	82.6	168	4	US-09-375-257-2	Sequence 2, Appli
16	113	81.9	23	1	US-08-333-565-10	Sequence 10, Appl
17	113	81.9	23	2	US-08-661-479-10	Sequence 10, Appl
18	102	73.9	59	2	US-08-733-505A-55	Sequence 55, Appl
19	102	73.9	59	2	US-08-733-505A-56	Sequence 56, Appl
20	102	73.9	59	2	US-08-733-505A-57	Sequence 57, Appl
21	102	73.9	59	2	US-08-733-505A-58	Sequence 58, Appl
22	86	62.3	16	1	US-08-333-565-26	Sequence 26, Appl
23	86	62.3	16	2	US-08-661-479-26	Sequence 26, Appl
24	61	44.2	11	2	US-08-733-505A-34	Sequence 34, Appl
25	61	44.2	11	2	US-08-706-741B-69	Sequence 69, Appl
26	61	44.2	11	2	US-08-924-695A-69	Sequence 69, Appl
27	51	37.0	66	2	US-08-867-087B-40	Sequence 40, Appl
28	48.5	35.1	904	4	US-09-328-352-4656	Sequence 4656, Ap
29	46	33.3	610	4	US-09-252-991A-19594	Sequence 19594, A
30	46	33.3	946	3	US-09-074-579-3	Sequence 3, Appli
31	46	33.3	946	3	US-09-388-774-3	Sequence 3, Appli
32	46	33.3	946	4	US-09-546-153-1	Sequence 1, Appli
33	45.5	33.0	906	4	US-09-252-991A-31458	Sequence 31458, A
34	45	32.6	229	4	US-09-252-991A-23807	Sequence 23807, A
35	45	32.6	303	4	US-09-328-352-5164	Sequence 5164, Ap
36	45	32.6	356	4	US-09-235-103-2	Sequence 2, Appli
37	45	32.6	356	4	US-09-235-103-4	Sequence 4, Appli
38	45	32.6	1064	4	US-09-252-991A-17508	Sequence 17508, A
39	44.5	32.2	903	4	US-09-252-991A-28775	Sequence 28775, A
40	44	31.9	125	4	US-09-328-352-7449	Sequence 7449, Ap
41	44	31.9	263	4	US-09-651-656-27	Sequence 27, Appl
42	44	31.9	263	4	US-09-650-855-27	Sequence 27, Appl
43	44	31.9	877	4	US-09-206-551-20	Sequence 20, Appl
44	44	31.9	1125	4	US-09-252-991A-18729	Sequence 18729, A
45	43.5	31.5	467	4	US-09-252-991A-18296	Sequence 18296, A

ALIGNMENTS

RESULT 1
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, Stanley J.
; TITLE OF INVENTION: Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
; TITLE OF INVENTION: REGULATOR
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,565
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15726A-000700
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..204
; OTHER INFORMATION: /note= "Deduced amino acid sequence
; OTHER INFORMATION: of mouse BAD."
US-08-333-565-2

Query Match 100.0%; Score 138; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 5.8e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 2
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             100.0%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 3
US-08-733-505A-1
; Sequence 1, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-733-505A-1

  Query Match             100.0%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 4
US-08-733-505A-12
; Sequence 12, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-12

  Query Match             100.0%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 5
US-08-733-505A-13
; Sequence 13, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-13

  Query Match             100.0%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 6
US-08-733-505A-14
; Sequence 14, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-14

  Query Match             100.0%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.8e-14;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 7
US-08-717-123-3
; Sequence 3, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
```

```
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-717-123-3

  Query Match             97.8%;  Score 135;  DB 2;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 1.7e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |||||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 8
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match             97.8%;  Score 135;  DB 4;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 1.7e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |||||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 9
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match             82.6%;  Score 114;  DB 1;  Length 166;
  Best Local Similarity   91.7%;  Pred. No. 2.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db        101 NLWAAQRYGRELRRMSDEFVDSFK 124


RESULT 10
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match             82.6%;  Score 114;  DB 2;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126
```

```
RESULT 11
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match             82.6%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 12
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match             82.6%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 13
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match             82.6%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 14
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match             82.6%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 15
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match             82.6%;  Score 114;  DB 4;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 2.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


Search completed: September 15, 2003, 17:45:05
Job time : 14.5571 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 20.6143 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	135	97.8	204	9	US-09-922-378-3	Sequence 3, Appli
2	135	97.8	204	14	US-10-066-179-3	Sequence 3, Appli
3	114	82.6	25	15	US-10-059-261-258	Sequence 258, App
4	114	82.6	168	9	US-09-922-378-2	Sequence 2, Appli
5	114	82.6	168	9	US-09-894-657-1	Sequence 1, Appli
6	114	82.6	168	9	US-09-894-657-7	Sequence 7, Appli
7	114	82.6	168	14	US-10-066-179-2	Sequence 2, Appli
8	71	51.4	15	15	US-10-174-105A-147	Sequence 147, App
9	50	36.2	215	15	US-10-156-761-9145	Sequence 9145, Ap
10	47	34.1	35	15	US-10-092-750-1	Sequence 1, Appli
11	47	34.1	138	15	US-10-092-750-241	Sequence 241, App
12	46	33.3	682	12	US-10-238-075-1077	Sequence 1077, Ap
13	46	33.3	946	9	US-09-828-423-3	Sequence 3, Appli
14	44	31.9	272	15	US-10-156-761-11541	Sequence 11541, A
15	44	31.9	426	9	US-09-815-242-5704	Sequence 5704, Ap
16	44	31.9	699	14	US-10-008-355-8	Sequence 8, Appli
17	44	31.9	705	9	US-09-815-242-12463	Sequence 12463, A
18	44	31.9	712	14	US-10-008-355-2	Sequence 2, Appli
19	44	31.9	877	12	US-10-369-294-20	Sequence 20, Appl
20	43	31.2	213	9	US-09-843-846-18	Sequence 18, Appl
21	43	31.2	232	10	US-09-881-752A-238	Sequence 238, App
22	43	31.2	270	11	US-09-934-455-162	Sequence 162, App
23	43	31.2	380	9	US-09-149-045-2	Sequence 2, Appli
24	43	31.2	380	15	US-10-166-359-2	Sequence 2, Appli
25	43	31.2	380	15	US-10-166-113-2	Sequence 2, Appli
26	43	31.2	380	15	US-10-166-357-2	Sequence 2, Appli
27	43	31.2	380	15	US-10-166-372-2	Sequence 2, Appli
28	43	31.2	380	15	US-10-184-722-3	Sequence 3, Appli
29	43	31.2	380	15	US-10-251-385-62	Sequence 62, Appl
30	43	31.2	380	15	US-10-251-385-198	Sequence 198, App
31	43	31.2	380	15	US-10-225-567A-233	Sequence 233, App
32	43	31.2	543	15	US-10-156-761-13485	Sequence 13485, A
33	43	31.2	571	9	US-09-815-242-11813	Sequence 11813, A
34	43	31.2	582	10	US-09-331-631A-22	Sequence 22, Appl
35	43	31.2	640	9	US-09-989-722-501	Sequence 501, App
36	43	31.2	640	9	US-09-989-723-501	Sequence 501, App
37	43	31.2	640	9	US-09-989-279-501	Sequence 501, App
38	43	31.2	640	9	US-09-989-727-501	Sequence 501, App
39	43	31.2	640	10	US-09-989-731-501	Sequence 501, App
40	43	31.2	640	10	US-09-989-732-501	Sequence 501, App
41	43	31.2	640	10	US-09-991-073-501	Sequence 501, App
42	43	31.2	640	10	US-09-909-320-292	Sequence 292, App
43	43	31.2	640	10	US-09-990-442-501	Sequence 501, App
44	43	31.2	640	10	US-09-991-163-501	Sequence 501, App
45	43	31.2	640	10	US-09-993-604-501	Sequence 501, App

ALIGNMENTS

RESULT 1
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE
; FILE REFERENCE: 480140.428D3
; CURRENT APPLICATION NUMBER: US/09/922,378
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-922-378-3

Query Match 97.8%; Score 135; DB 9; Length 204;
Best Local Similarity 96.2%; Pred. No. 3.3e-12;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||:||||||||||
Db 140 NLWAAQRYGRELRRMTDEFEGSFKGL 165

RESULT 2
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE

```
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             97.8%;  Score 135;  DB 14;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 3.3e-12;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 3
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match             82.6%;  Score 114;  DB 15;  Length 25;
  Best Local Similarity   91.7%;  Pred. No. 4.8e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db          1 NLWAAQRYGRELRRMSDEFVDSFK 24


RESULT 4
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match             82.6%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 5
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: SYNORAB01
;            CLONE: 358673
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             82.6%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 6
US-09-894-657-7
```

```
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: GenBank
;            CLONE: 1683637
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match             82.6%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 7
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match             82.6%;  Score 114;  DB 14;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 3.4e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECI
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHI
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match             51.4%;  Score 71;  DB 15;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00061;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFEGS 22
              ||||||||||||||
Db          1 GRELRRMSDEFEGS 14


RESULT 9
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
```

```
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

  Query Match             36.2%;  Score 50;  DB 15;  Length 215;
  Best Local Similarity   56.2%;  Pred. No. 12;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 QRYGRELRRMSDEFEG 21
              :|:| :|||| || :|
Db        108 ERWGGDLRRMRDEADG 123


RESULT 10
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-1

  Query Match             34.1%;  Score 47;  DB 15;  Length 35;
  Best Local Similarity   45.5%;  Pred. No. 5;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              :| ||    ||||: |||   :
Db         15 IWIAQ----ELRRIGDEFNAYY 32


RESULT 11
US-10-092-750-241
; Sequence 241, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-241

  Query Match             34.1%;  Score 47;  DB 15;  Length 138;
  Best Local Similarity   45.5%;  Pred. No. 20;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              :| ||    ||||: |||   :
Db         86 IWIAQ----ELRRIGDEFNAYY 103


RESULT 12
US-10-238-075-1077
; Sequence 1077, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
;  APPLICANT: I.N.S.E.R.M.
;  TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isc
;  TITLE OF INVENTION:  E.coli, and biological uses of these polynucleotides and of t
;  FILE REFERENCE: BLANDINE
;  CURRENT APPLICATION NUMBER: US/10/238,075
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 0003145
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 1576
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1077
;   LENGTH: 682
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-238-075-1077

  Query Match             33.3%;  Score 46;  DB 12;  Length 682;
  Best Local Similarity   40.0%;  Pred. No. 1.5e+02;
  Matches   10;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 LWAAQRYGRELRRMSDEFEGSFKGL 26
              :||||| | :: |: : |     ||
Db        610 IWAAQRNGAKVPRVRNGFTSMDIGL 634


RESULT 13
US-09-828-423-3
; Sequence 3, Application US/09828423
; Patent No. US20020099178A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Guegler, Karl J.
;                   Patterson, Chandra
;        TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
;                            INHIBITOR HEAVY CHAIN PRECURSOR
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/828,423
;             FILING DATE: 05-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/388,774
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER: PF-0505 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
```

```
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 946 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GENEBANK
;             CLONE: gi33985
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-09-828-423-3

  Query Match             33.3%;  Score 46;  DB 9;  Length 946;
  Best Local Similarity   30.8%;  Pred. No. 2.1e+02;
  Matches    8;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ::|  :  |     : | ||| | |:
Db        212 DVWVIEPQGLRFLHVPDTFEGHFDGV 237


RESULT 14
US-10-156-761-11541
; Sequence 11541, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11541
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11541

  Query Match             31.9%;  Score 44;  DB 15;  Length 272;
  Best Local Similarity   53.3%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 WAAQRYGRELRRMSD 17
              | |  :| |||| :|
Db         29 WIAAAHGAELRRAAD 43


RESULT 15
US-09-815-242-5704
; Sequence 5704, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5704
;   LENGTH: 426
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-5704

  Query Match             31.9%;  Score 44;  DB 9;  Length 426;
  Best Local Similarity   36.8%;  Pred. No. 1.8e+02;
  Matches    7;  Conservative    4;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          8 YGRELRRMSDEFEGSFKGL 26
              :|  |||  : |:|  | :
Db        386 FGGSLRRQDENFDGKIKAI 404


Search completed: September 15, 2003, 17:47:52
Job time : 21.6143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 11.7 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	204	2	A55671	bad protein - mous
2	54	39.1	946	2	JC5575	inter-alpha-trypsi
3	53	38.4	223	2	D70760	hypothetical prote
4	53	38.4	946	2	S54354	inter-alpha-inhibi
5	52	37.7	370	2	S38185	2-dehydro-3-deoxy-
6	51	37.0	232	2	A42095	floral homeotic pr
7	50	36.2	374	2	C84338	spermidine/putresc
8	50	36.2	516	2	A96753	probable threonine
9	49	35.5	453	2	E83517	conserved hypothet
10	48.5	35.1	134	2	S40376	Ig kappa chain - h
11	48.5	35.1	314	2	T02975	annexin P35 - maiz
12	48	34.8	206	2	C36365	transforming prote
13	48	34.8	220	2	F72289	oxidoreductase, so
14	48	34.8	526	2	T08545	threonine synthase
15	47	34.1	597	2	G82308	oxaloacetate decar
16	47	34.1	967	2	F82668	oxoglutarate dehyd
17	47	34.1	5138	2	B96695	hypothetical prote
18	46.5	33.7	314	2	T02961	annexin P33 - maiz
19	46.5	33.7	435	2	A44308	Antho-RFamide prec
20	46.5	33.7	1140	2	T09486	hypothetical prote
21	46	33.3	399	2	T35440	probable polyamine
22	46	33.3	946	1	IYHU2	inter-alpha-trypsi
23	46	33.3	1164	2	T24806	hypothetical prote
24	46	33.3	1378	2	A81393	DNA-directed RNA p
25	45.5	33.0	261	2	G69510	conserved hypothet
26	45.5	33.0	287	2	S43852	neuropeptide Pol-R
27	45.5	33.0	334	2	A39172	Antho-RFamide neur
28	45.5	33.0	562	2	C71473	hypothetical prote
29	45.5	33.0	905	2	G83314	NADH dehydrogenase
30	45	32.6	165	2	S59899	chlorocruorin chai
31	45	32.6	295	2	F83201	conserved hypothet
32	45	32.6	346	2	H95406	conserved hypothet
33	45	32.6	591	2	B44465	sodium ion pump ox
34	45	32.6	591	2	AB0509	oxaloacetate decar
35	45	32.6	591	2	AE0909	oxaloacetate decar
36	45	32.6	596	2	A28088	oxaloacetate decar
37	45	32.6	715	2	S52675	probable membrane
38	45	32.6	864	1	VCLJG4	env polyprotein -
39	45	32.6	1263	2	T19472	hypothetical prote
40	45	32.6	1557	2	T28811	hypothetical prote
41	45	32.6	2325	2	T15566	hypothetical prote
42	44.5	32.2	75	2	T01993	hypothetical prote
43	44.5	32.2	455	2	D83264	hypothetical prote
44	44.5	32.2	536	2	AG1482	hypothetical prote
45	44.5	32.2	910	2	G91024	NADH dehydrogenase

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 100.0%; Score 138; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165
```

RESULT 2
JC5575
inter-alpha-trypsin inhibitor heavy chain 2 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
C;Accession: JC5575; PC4485
R;Nakatani, T.; Suzuki, Y.; Yamamoto, T.; Sinohara, H.
J. Biochem. 122, 71-82, 1997
A;Title: Molecular cloning and sequencing of cDNAs encoding three heavy-chain precurs
sin inhibitor heavy chain family.
A;Reference number: JC5574; MUID:97420688; PMID:9276673
A;Accession: JC5575
A;Molecule type: mRNA
A;Residues: 1-946 <NAK>
A;Cross-references: DDBJ:D89286; NID:g1694689; PIDN:BAA13939.1; PID:g1694690
A;Experimental source: liver
A;Accession: PC4485
A;Molecule type: protein
A;Residues: 55-64;140-146;151-156;424-447;500-528;577-605 <NA2>
C;Comment: In the plasma three inter-alpha-trypsin inhibitor heavy chains 1, 2 and 3
that the complexes play important role for panceatic cancer.
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
F;261-264,717-916/Disulfide bonds: #status predicted

Query Match 39.1%; Score 54; DB 2; Length 946;
Best Local Similarity 34.6%; Pred. No. 8.8;

Matches 9; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGV 237
```

RESULT 3
D70760
hypothetical protein Rv2014 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70760
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70760
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <COL>
A;Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CAA98415.1; PID:e1299911;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2014

Query Match 38.4%; Score 53; DB 2; Length 223;
Best Local Similarity 58.8%; Pred. No. 2.8;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181
```

RESULT 4
S54354
inter-alpha-inhibitor H2 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999
C;Accession: S54354
R;Chan, P.; Risler, J.L.; Raguenez, G.; Salier, J.P.
Biochem. J. 306, 505-512, 1995
A;Title: The three heavy-chain precursors for the inter-alpha-inhibitor family in mouse:
A;Reference number: S54353; MUID:95194326; PMID:7534067
A;Accession: S54354
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-946 <CHA>
A;Cross-references: EMBL:X70392; NID:g695633; PIDN:CAA49842.1; PID:g695634
C;Superfamily: inter-alpha-trypsin inhibitor complex component II

Query Match 38.4%; Score 53; DB 2; Length 946;
Best Local Similarity 34.6%; Pred. No. 12;
Matches 9; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237
```

RESULT 5
S38185
2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) ARO4 - yeast (Saccharomyces ce
N;Alternate names: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DAHP synthase;
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C;Accession: S38185; S46126; S46130; JN0322; B48651
R;Doignon, F.; Biteau, N.; Aigle, M.; Crouzet, M.
Yeast 9, 1131-1137, 1993
A;Title: The complete sequence of a 6794 bp segment located on the right arm of chrom
A;Reference number: S38185; MUID:94078675; PMID:8256522
A;Accession: S38185
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <DOI>
A;Cross-references: GB:L20296; NID:g311101; PIDN:AAA65607.1; PID:g311102
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46126
A;Molecule type: DNA
A;Residues: 1-370 <ALJ>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46130
A;Molecule type: DNA
A;Residues: 1-370 <AIG>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR
R;Kuenzler, M.; Paravicini, G.; Egli, C.M.; Irniger, S.; Braus, G.H.
Gene 113, 67-74, 1992
A;Title: Cloning, primary structure and regulation of the ARO4 gene, encoding the tyr
A;Reference number: JN0322; MUID:92225349; PMID:1348717
A;Accession: JN0322
A;Molecule type: DNA
A;Residues: 1-204,208-370 <KUE>
A;Cross-references: EMBL:X61107
R;Kuenzler, M.; Balmelli, T.; Egli, C.M.; Paravicini, G.; Braus, G.H.
J. Bacteriol. 175, 5548-5558, 1993
A;Title: Cloning, primary structure, and regulation of the HIS7 gene encoding a bifur
A;Reference number: A48651; MUID:93374850; PMID:8366040
A;Accession: B48651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 352-370 <KU2>
A;Cross-references: GB:X61107
C;Comment: This enzyme catalyzes the condensation of phosphoenolpyruvate and D-erythr
C;Genetics:
A;Gene: SGD:ARO4
A;Cross-references: SGD:S0000453; MIPS:YBR249c
A;Map position: 2R
C;Function:
A;Description: aldehyde-lyase; carbon-carbon lyase
A;Pathway: aromatic amino acid biosynthesis; shikimate pathway
A;Note: first step in shikimate pathway
C;Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase
C;Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; cy

Query Match 37.7%; Score 52; DB 2; Length 370;
Best Local Similarity 47.6%; Pred. No. 6.7;
Matches 10; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEG 21
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100
```

RESULT 6
A42095
floral homeotic protein APETALA3 (AP3) - Arabidopsis thaliana
N;Alternate names: homeotic protein APETALA3; MADS-box regulatory protein AP3
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A42095; S52633; T47593
R;Jack, T.; Brockman, L.L.; Meyerowitz, E.M.
Cell 68, 683-697, 1992
A;Title: The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is
A;Reference number: A42095; MUID:92154682; PMID:1346756
A;Accession: A42095
A;Status: preliminary
A;Molecule type: mRNA

A;Residues: 1-232 <JAC>
A;Cross-references: GB:M86357; NID:g166607; PIDN:AAA32740.1; PID:g166608
A;Experimental source: petals, stamens
A;Note: sequence extracted from NCBI backbone (NCBIN:82520, NCBIP:82521)
R;Okamoto, H.; Yano, A.; Shiraishi, H.; Okada, K.; Shimura, Y.
Plant Mol. Biol. 26, 465-472, 1994
A;Title: Genetic complementation of a floral homeotic mutation, apetala3, with an Arabid
A;Reference number: S52633; MUID:95036018; PMID:7948893
A;Accession: S52633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <OKA>
A;Cross-references: GB:D21125
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24469
A;Accession: T47593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <BLO>
A;Cross-references: EMBL:AL132971
A;Experimental source: cultivar Columbia; BAC clone T12E18
C;Genetics:
A;Map position: 3
A;Introns: 63/2; 85/3; 106/2; 139/3; 153/3; 168/3
A;Note: T12E18.30
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match 37.0%; Score 51; DB 2; Length 232;
  Best Local Similarity 44.4%; Pred. No. 5.8;
  Matches 12; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

Qy 6 QRYG--------RELRRMSDEFEGSFK 24
Db 107 QRLGECLDELDIQELRRLEDEMENTFK 133

RESULT 7
C84338
spermidine/putrescine ABC transporter [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84338
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84338
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <STO>
A;Cross-references: GB:AE004437; NID:g10581314; PIDN:AAG20071.1; GSPDB:GN00138
C;Genetics:
A;Gene: potA2

  Query Match 36.2%; Score 50; DB 2; Length 374;
  Best Local Similarity 76.9%; Pred. No. 13;
  Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 11 ELRRMSDEFEGSF 23
Db 197 ELRRLSDAVEGSF 209

RESULT 8
A96753
probable threonine synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96753
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE005173; NID:g5903070; PIDN:AAD55628.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3N23.1
A;Map position: 1

  Query Match 36.2%; Score 50; DB 2; Length 516;
  Best Local Similarity 35.3%; Pred. No. 18;
  Matches 12; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

Qy 1 NLWAAQRYGRELRRMSD--------EFEGSFKGL 26
Db 163 NLFWAERFGKQYLQMNDLWVKHCGISHTGSFKDL 196

RESULT 9
E83517
conserved hypothetical protein PA1031 [imported] - Pseudomonas aeruginosa (strain PAO
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83517
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <STO>
A;Cross-references: GB:AE004535; GB:AE004091; NID:g9946936; PIDN:AAG04420.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1031

  Query Match 35.5%; Score 49; DB 2; Length 453;
  Best Local Similarity 55.6%; Pred. No. 23;
  Matches 10; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 3 WAAQRYGR--ELRRMSDE 18
Db 65 WASERQGREEELRRLASE 82

RESULT 10
S40376
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40376
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40376
A;Status: preliminary; translation not shown

```
A;Molecule type: mRNA
A;Residues: 1-134 <KLE>
A;Cross-references: EMBL:X72486; NID:g441440; PIDN:CAA51154.1; PID:g441441
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-113/Domain: immunoglobulin homology <IMM>

  Query Match             35.1%;  Score 48.5;  DB 2;  Length 134;
  Best Local Similarity   38.2%;  Pred. No. 7.8;
  Matches   13;  Conservative    1;  Mismatches    9;  Indels   11;  Gaps    1;

Qy          3 WAAQRYGRELRRM----------SDEFEGSFKG 25
              |  || ||  ||:           || | ||  |
Db         58 WFRQRPGRSPRRLIYNVSKRDSGVSDRFSGSGSG 91


RESULT 11
T02975
annexin P35 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02975
R;Battey, N.H.; James, N.C.; Greenland, A.J.
Plant Physiol. 112, 1391-1396, 1996
A;Title: CDNA isolation and gene expression of maize annexins P33 and P35.
A;Reference number: Z14796; MUID:97092863; PMID:8938425
A;Accession: T02975
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-314 <BAT>
A;Cross-references: EMBL:X98245; NID:g1370602; PIDN:CAA66901.1; PID:g1370603
A;Experimental source: cultivar clipper; root tip
C;Superfamily: annexin I; annexin repeat homology
F;14-85/Domain: annexin repeat homology <AXR>

  Query Match             35.1%;  Score 48.5;  DB 2;  Length 314;
  Best Local Similarity   47.6%;  Pred. No. 19;
  Matches   10;  Conservative    4;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          5 AQRYGRE-LRRMSDEFEGSFK 24
              |: ||:| || : ||  | |:
Db         54 AEAYGKELLRALGDEIHGKFE 74


RESULT 12
C36365
transforming protein homolog MRAS3 - Rhizomucor racemosus
C;Species: Rhizomucor racemosus
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 19-Jan-2001
C;Accession: C36365
R;Casale, W.L.; McConnell, D.G.; Wang, S.Y.; Lee, Y.J.; Linz, J.E.
Mol. Cell. Biol. 10, 6654-6663, 1990
A;Title: Expression of a gene family in the dimorphic fungus Mucor racemosus which exhi
A;Reference number: A36365; MUID:91061774; PMID:1701021
A;Accession: C36365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <CAS>
A;Cross-references: GB:M55177
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;11-126/Domain: translation elongation factor Tu homology <ETU>
F;17-24/Region: nucleotide-binding motif A (P-loop)
F;123-126/Region: GTP-binding NKXD motif
F;153-155/Region: GTP-binding SAK/L motif
F;23,24,42,123,124,126,153/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             34.8%;  Score 48;  DB 2;  Length 206;
  Best Local Similarity   62.5%;  Pred. No. 14;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 RELRRMSDEFEGSFKG 25
              ||:|||: | ||  ||
Db        169 REIRRMNKEQEGRSKG 184


RESULT 13
F72289
oxidoreductase, sol/devB family - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72289
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <ARN>
A;Cross-references: GB:AE001772; GB:AE000512; NID:g4981693; PIDN:AAD36230.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1154
C;Superfamily: yeast SOL3 protein

  Query Match             34.8%;  Score 48;  DB 2;  Length 220;
  Best Local Similarity   34.8%;  Pred. No. 15;
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 AAQRYGRELRRMSDEFEGSFKGL 26
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133


RESULT 14
T08545
threonine synthase (EC 4.2.3.1) precursor - Arabidopsis thaliana
N;Alternate names: protein F27B13.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C;Accession: T08545; S71362; S74307
R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.;
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16442
A;Accession: T08545
A;Molecule type: DNA
A;Residues: 1-526 <BEV>
A;Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.80
A;Experimental source: cultivar Columbia; BAC clone F27B13
R;Curien, G.; Dumas, R.; Ravanel, S.; Douce, R.
FEBS Lett. 390, 85-90, 1996
A;Title: Characterization of an Arabidopsis thaliana cDNA encoding an S-adenosylmethi
A;Reference number: S71362; MUID:96314555; PMID:8706836
A;Accession: S71362
A;Molecule type: mRNA
A;Residues: 'L',3-526 <CUR>
A;Cross-references: EMBL:L41666; NID:g1448916; PIDN:AAB04607.1; PID:g1448917
A;Accession: S74307
A;Molecule type: protein
A;Residues: 40-54 <CUI>
C;Genetics:
A;Gene: ATSP:F27B13.80
A;Map position: 4
A;Genome: nuclear
C;Keywords: carbon-oxygen lyase; chloroplast
F;1-39/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;40-526/Product: threonine synthase #status experimental <MAT>

  Query Match             34.8%;  Score 48;  DB 2;  Length 526;
  Best Local Similarity   35.3%;  Pred. No. 37;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    8;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSD--------EFEGSFKGL 26
```

```
              ||: |:|:|::   |:|          |||| |
Db        172 NLFWAERFGKQFLGMNDLWVKHCGISHTGSFKDL 205


RESULT 15
G82308
oxaloacetate decarboxylase, alpha chain VC0550 [similarity] - Vibrio cholerae (strain N1
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001
C;Accession: G82308
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: G82308
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-597 <HEI>
A;Cross-references: GB:AE004141; GB:AE003852; NID:g9654976; PIDN:AAF93718.1; GSPDB:GN001
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VC0550
A;Map position: 1
C;Superfamily: Klebsiella pneumoniae oxaloacetate decarboxylase alpha chain; lipoyl/biot

  Query Match             34.1%;  Score 47;  DB 2;  Length 597;
  Best Local Similarity   47.4%;  Pred. No. 60;
  Matches    9;  Conservative    4;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          8 YGRELRRMSDEFEGSFKGL 26
              | ||:|:   :|||  ||:
Db        272 YFREVRKKYAKFEGQLKGV 290


Search completed: September 15, 2003, 17:27:00
Job time : 12.7 secs
```

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 6.12857 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	100.0	204	1	BAD_MOUSE	Q61337 mus musculu
2	138	100.0	205	1	BAD_RAT	O35147 rattus norv
3	114	82.6	168	1	BAD_HUMAN	Q92934 homo sapien
4	54	39.1	946	1	ITH2_MESAU	P97279 mesocricetu
5	53	38.4	946	1	ITH2_MOUSE	Q61703 mus musculu
6	52	37.7	370	1	AROG_YEAST	P32449 saccharomyc
7	51	37.0	232	1	AP3_ARATH	P35632 arabidopsis
8	50	36.2	851	1	CE05_MOUSE	Q8k2h3 mus musculu
9	49.5	35.9	506	1	MATK_LEDPA	O62992 ledum palus
10	49.5	35.9	506	1	MATK_RHOFR	O62984 rhododendro
11	49.5	35.9	506	1	MATK_RHOTS	O62991 rhododendro
12	49	35.5	453	1	RMUC_PSEAE	Q9i4u3 pseudomonas
13	48	34.8	205	1	RAS3_RHIRA	P22280 rhizomucor
14	48	34.8	220	1	6PGL_THEMA	Q9x0n8 thermotoga
15	48	34.8	519	1	THRC_SOLTU	Q9mt28 solanum tub
16	48	34.8	526	1	THRC_ARATH	Q9s7b5 arabidopsis
17	47	34.1	198	1	BIM_HUMAN	O43521 homo sapien
18	46.5	33.7	429	1	FMR2_ANTEL	Q16994 anthopleura
19	46.5	33.7	435	1	FMR1_ANTEL	P10419 anthopleura
20	46	33.3	946	1	ITH2_HUMAN	P19823 homo sapien
21	46	33.3	1378	1	RPOB_CAMJE	Q46124 campylobact
22	45.5	33.0	287	1	PRFA_POLPE	P21259 polyorchis
23	45.5	33.0	334	1	FMRA_CALPA	Q01133 calliactis
24	45.5	33.0	507	1	MATK_LOIPR	O47169 loiseleuria
25	45	32.6	328	1	SNF4_KLULA	Q9p869 kluyveromyc
26	45	32.6	590	1	DCOA_SALTY	Q03030 salmonella
27	45	32.6	595	1	DCOA_KLEPN	P13187 klebsiella
28	45	32.6	653	1	HT2A_HUMAN	Q13049 homo sapien
29	45	32.6	865	1	ENV_SIVAT	P05886 simian immu
30	45	32.6	1535	1	LML1_CAEEL	Q18823 caenorhabdi
31	44.5	32.2	907	1	NUOG_ECOLI	P33602 escherichia
32	44.5	32.2	907	1	NUOG_SALTY	P33900 salmonella
33	44	31.9	196	1	BIM_MOUSE	O54918 mus musculu
34	44	31.9	196	1	BIM_RAT	O88498 rattus norv
35	44	31.9	262	1	END8_ECO57	Q8x9c6 escherichia
36	44	31.9	262	1	END8_ECOLI	P50465 escherichia
37	44	31.9	262	1	END8_SALTI	Q8z8d2 salmonella
38	44	31.9	262	1	END8_SALTY	Q8zqu6 salmonella
39	44	31.9	768	1	ENV_SIVA1	P27757 simian immu
40	44	31.9	877	1	ENV_SIVAG	P27977 simian immu
41	44	31.9	915	1	CE05_HUMAN	Q9nyf5 homo sapien
42	44	31.9	5596	1	MDN1_HUMAN	Q9nu22 homo sapien
43	44	31.9	8797	1	SNE1_HUMAN	Q8nf91 homo sapien
44	43.5	31.5	217	1	UREF_SYNY3	P73327 synechocyst
45	43.5	31.5	1014	1	UVRA_STRCO	Q9z507 streptomyce

ALIGNMENTS

RESULT 1
BAD_MOUSE
ID BAD_MOUSE STANDARD; PRT; 204 AA.
AC Q61337;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE 6) (Bcl-xL/Bcl-2 associated death promoter).
GN BAD OR BBC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Thymus;
RX MEDLINE=95136361; PubMed=7834748;
RA Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT promotes cell death.";
RL Cell 80:285-291(1995).
RN [2]
RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX MEDLINE=98022383; PubMed=9381178;
RA Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT "Interleukin-3-induced phosphorylation of BAD through the protein
RT kinase Akt.";
RL Science 278:687-689(1997).
RN [3]
RP MUTAGENESIS OF SERINE RESIDUES.
RX MEDLINE=20403302; PubMed=10949026;
RA Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA Greenberg M.E.;
RT "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT BH3 domain phosphorylation.";
RL Mol. Cell 6:41-51(2000).
CC -!- FUNCTION: Promotes cell death. Successfully competes for the
CC binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC of heterodimerization of these proteins with BAX. Can reverse the
CC death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC Appears to act as a link between growth factor receptor signaling
CC and the apoptotic pathways.
CC -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC phosphorylation, locates to the cytoplasm.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC with 14-3-3 proteins. This interaction then facilitates the
CC phosphorylation at Ser-155, a site within the BH3 domain, leading
CC to the release of Bcl-x(L) and the promotion of cell survival.

```
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             100.0%;  Score 138;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 2
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;
```

```
  Query Match           100.0%;  Score 138;  DB 1;  Length 205;
  Best Local Similarity  100.0%;  Pred. No. 1.6e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db        141 NLWAAQRYGRELRRMSDEFEGSFKGL 166


RESULT 3
BAD_HUMAN
ID   BAD_HUMAN      STANDARD;      PRT;   168 AA.
AC   Q92934; O14803;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE   XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN   BAD OR BBC6 OR BCL2L8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT   "A human protein that interacts with Bcl-2 and have homology to mouse
RT   BAD.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX   MEDLINE=97083574; PubMed=8929532;
RA   Wang H.-G., Rapp U.R., Reed J.C.;
RT   "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL   Cell 87:629-638(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Takayama S., Reed J.C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND DIMERIZATION.
RC   TISSUE=Bone marrow;
RX   MEDLINE=98049554; PubMed=9388232;
RA   Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA   Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT   "Dimerization properties of human BAD.";
RL   J. Biol. Chem. 272:30866-30872(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match           82.6%;  Score 114;  DB 1;  Length 168;
```

```
  Best Local Similarity   91.7%;  Pred. No. 5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 4
ITH2_MESAU
ID   ITH2_MESAU     STANDARD;      PRT;   946 AA.
AC   P97279;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2) (HC2).
GN   ITIH2.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97420688; PubMed=9276673;
RA   Nakatani T., Suzuki Y., Yamamoto T., Sinohara H.;
RT   "Molecular cloning and sequencing of cDNAs encoding three heavy-chain
RT   precursors of the inter-alpha-trypsin inhibitor in Syrian hamster:
RT   implications for the evolution of the inter-alpha-trypsin inhibitor
RT   heavy chain family.";
RL   J. Biochem. 122:71-82(1997).
RN   [2]
RP   SEQUENCE OF 55-64; 140-146; 151-156; 424-447; 500-528 AND 577-605,
RP   AND SUBUNITS.
RC   TISSUE=Plasma;
RX   MEDLINE=97018241; PubMed=8864857;
RA   Yamamoto T., Yamamoto K., Sinohara H.;
RT   "Inter-alpha-trypsin inhibitor and its related proteins in Syrian
RT   hamster urine and plasma.";
RL   J. Biochem. 120:145-152(1996).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89286; BAA13939.1; -.
DR   PIR; JC5575; JC5575.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                 H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                 (BY SIMILARITY).
FT   CONFLICT    510    510       V -> Y (IN REF. 2).
FT   CONFLICT    595    595       E -> I (IN REF. 2).
SQ   SEQUENCE   946 AA;  106580 MW;  CA8BF565458E7B2E CRC64;

  Query Match             39.1%;  Score 54;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 3.4;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGV 237


RESULT 5
ITH2_MOUSE
ID   ITH2_MOUSE     STANDARD;      PRT;   946 AA.
AC   Q61703;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2).
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6N; TISSUE=Liver;
RX   MEDLINE=95194326; PubMed=7534067;
RA   Chan P., Risler J.-L., Raguenez G., Salier J.-P.;
RT   "The three heavy-chain precursors for the inter-alpha-inhibitor
RT   family in mouse: new members of the multicopper oxidase protein group
RT   with differential transcription in liver and brain.";
RL   Biochem. J. 306:505-512(1995).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH LIVER AND BRAIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X70392; CAA49842.1; -.
DR   PIR; S54354; S54354.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                (BY SIMILARITY).
SQ   SEQUENCE   946 AA;  105927 MW;  40DB6716433ED9DC CRC64;

  Query Match             38.4%;  Score 53;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 4.8;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237


RESULT 6
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match             37.7%;  Score 52;  DB 1;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 2.4;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEG 21
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 7
AP3_ARATH
ID   AP3_ARATH      STANDARD;      PRT;   232 AA.
AC   P35632; Q39003; Q8LB79; Q9S7Q3; Q9SQ14; Q9SQ15; Q9SQ16; Q9SQ17;
AC   Q9SQ18; Q9SQ19; Q9SQ20; Q9SQ21; Q9SQ22; Q9SX13;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Floral homeotic protein APETALA3.
GN   AP3 OR AT3G54340 OR T12E18_30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Petal;
RX   MEDLINE=92154682; PubMed=1346756;
RA   Jack T., Brockman L.L., Meyerowitz E.M.;
RT   "The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS
RT   box and is expressed in petals and stamens.";
RL   Cell 68:683-697(1992).
```

```
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=95036018; PubMed=7948893;
RA   Okamoto H., Yano A., Shiraishi H., Okada K., Shimura Y.;
RT   "Genetic complementation of a floral homeotic mutation, apetala3,
RT   with an Arabidopsis thaliana gene homologous to DEFICIENS of
RT   Antirrhinum majus.";
RL   Plant Mol. Biol. 26:465-472(1994).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   STRAIN=cv. Bla-1, cv. Bretagny, cv. Bs-1, cv. Bu-0, cv. Bu-2,
RC   cv. Chi-1, cv. Co-1, cv. Columbia, cv. Corsacalla-1, cv. Cvi-0,
RC   cv. Gr-3, cv. Jl-1, cv. Kas-1, cv. Kent, cv. Landsberg erecta,
RC   cv. Li-3, cv. Li-8, and cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci: departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 36-128 FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=99311297; PubMed=10382288;
RA   Brunel D., Froger N., Pelletier G.;
RT   "Development of amplified consensus genetic markers (ACGM) in Brassica
RT   napus from Arabidopsis thaliana sequences of known biological
RT   function.";
RL   Genome 42:387-402(1999).
RN   [8]
RP   FUNCTION.
RX   PubMed=8565821;
RA   Krizek B.A., Meyerowitz E.M.;
RT   "The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient
RT   to provide the B class organ identity function.";
RL   Development 122:11-22(1996).
RN   [9]
RP   CHARACTERIZATION.
RX   PubMed=8643482;
RA   Riechmann J.L., Krizek B.A., Meyerowitz E.M.;
RT   "Dimerization specificity of Arabidopsis MADS domain homeotic proteins
RT   APETALA1, APETALA3, PISTILLATA, and AGAMOUS.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:4793-4798(1996).
RN   [10]
RP   GENETIC REGULATION.
RX   PubMed=11283333;
RA   Ng M., Yanofsky M.F.;
RT   "Activation of the Arabidopsis B class homeotic genes by APETALA1.";
RL   Plant Cell 13:739-753(2001).
RN   [11]
RP   CHARACTERIZATION.
RX   PubMed=11206550;
RA   Honma T., Goto K.;
RT   "Complexes of MADS-box proteins are sufficient to convert leaves into
RT   floral organs.";
RL   Nature 409:525-529(2001).
CC   -!- FUNCTION: Probable transcription factor involved in the genetic
CC       control of flower development. Is required for normal development
CC       of petals and stamens in the wild-type flower. Forms an
CC       heterodimer with PISTILLATA that is required for autoregulation of
CC       both AP3 and PI genes. AP3/PI heterodimer interacts with APETALA1
CC       or SEPALLATA3 to form a ternary complex that could be responsible
CC       for the regulation of the genes involved in the flower
CC       development.
CC   -!- SUBUNIT: Forms an heterodimer with PISTILLATA, capable of binding
CC       to CArG-box sequences. AP3/PI heterodimer binds AP1 or SEP3 to
CC       form complexes.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Expressed in petals and stamens.
CC   -!- INDUCTION: Positively regulated by the meristem identity proteins
CC       APETALA1 and LEAFY with the cooperation of UFO.
CC   -!- MISCELLANEOUS: Mutations in AP3 cause transformation of petals
CC       into sepals and stamina into carpels.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: Contains 1 K-box dimerization domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M86357; AAA32740.1; -.
DR   EMBL; D21125; BAA04665.1; -.
DR   EMBL; AF115798; AAD51887.1; -.
DR   EMBL; AF115799; AAD51888.1; -.
DR   EMBL; AF115800; AAD51889.1; -.
DR   EMBL; AF115801; AAD51890.1; -.
DR   EMBL; AF115802; AAD51891.1; -.
DR   EMBL; AF115803; AAD51892.1; -.
DR   EMBL; AF115804; AAD51893.1; -.
DR   EMBL; AF115805; AAD51894.1; -.
DR   EMBL; AF115806; AAD51895.1; -.
DR   EMBL; AF115807; AAD51896.1; -.
DR   EMBL; AF115808; AAD51897.1; -.
DR   EMBL; AF115809; AAD51898.1; -.
```

```
DR   EMBL; AF115810; AAD51899.1; -.
DR   EMBL; AF115811; AAD51900.1; -.
DR   EMBL; AF115812; AAD51901.1; -.
DR   EMBL; AF115813; AAD51902.1; -.
DR   EMBL; AF115814; AAD51903.1; -.
DR   EMBL; AL132971; CAB81799.1; -.
DR   EMBL; AY087369; AAM64919.1; -.
DR   EMBL; AY070397; AAL49893.1; -.
DR   EMBL; AY142590; AAN13159.1; -.
DR   EMBL; AF056541; AAD41557.1; -.
DR   PIR; A42095; A42095.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T01776; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Flowering; Transcription regulation; Activator; Developmental protein;
KW   Nuclear protein; DNA-binding; Coiled coil; Polymorphism.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       93    165       K-BOX.
FT   DOMAIN       75    164       COILED COIL (POTENTIAL).
FT   VARIANT      31     31       K -> R (in strain cv. Lisse).
FT   VARIANT      47     47       M -> T (in strain cv. Bretagny).
FT   VARIANT      61     61       N -> D (in strain cv. Corsacalla-1).
FT   VARIANT      73     73       T -> S (in strain cv. Li-8).
FT   VARIANT     109    109       L -> V (in strain cv. Kas-1).
FT   VARIANT     115    115       E -> K (in strains cv. Chi-1 and cv.
FT                                Gr-3).

  Query Match             37.0%;  Score 51;  DB 1;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 2.1;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 8
CE05_MOUSE
ID   CE05_MOUSE     STANDARD;      PRT;   851 AA.
AC   Q8K2H3;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Protein C5orf5 homolog.
GN   C5ORF5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Belongs to the FAM13 family.
CC   -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC031465; AAH31465.1; -.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   GTPase activation.
FT   DOMAIN       23    212       RHO-GAP.
FT   DOMAIN      189    256       GLU-RICH.
SQ   SEQUENCE   851 AA;  97054 MW;  C2B26669FB6DB2CE CRC64;

  Query Match             36.2%;  Score 50;  DB 1;  Length 851;
  Best Local Similarity   45.5%;  Pred. No. 12;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              || |:   ::||:|  ||| :|
Db        784 LWKARAEKKKLRKMLREFEEAF 805


RESULT 9
MATK_LEDPA
ID   MATK_LEDPA     STANDARD;      PRT;   506 AA.
AC   O62992;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Ledum palustre (Wild rosemary).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75583;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB012751; BAA25872.1; -.
```

```
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60412 MW;  CFEA926307DAC85E CRC64;

  Query Match             35.9%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 8.2;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 10
MATK_RHOFR
ID   MATK_RHOFR     STANDARD;      PRT;   506 AA.
AC   O62984;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron ferrugineum (Alpenrose).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49622;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB012741; BAA25862.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60534 MW;  ADA44B25E92436E8 CRC64;

  Query Match             35.9%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 8.2;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 11
MATK_RHOTS
ID   MATK_RHOTS     STANDARD;      PRT;   506 AA.
AC   O62991;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron tsusiophyllum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49629;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AB012750; BAA25871.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60569 MW;  AEE12FF8809C223E CRC64;

  Query Match             35.9%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 8.2;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 12
RMUC_PSEAE
ID   RMUC_PSEAE     STANDARD;      PRT;   453 AA.
AC   Q9I4U3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA recombination protein rmuC homolog.
GN   RMUC OR PA1031.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
```

```
CC   -!- FUNCTION: Involved in DNA recombination (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RMUC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004535; AAG04420.1; -.
DR   PIR; E83517; E83517.
DR   InterPro; IPR003798; DUF195.
DR   Pfam; PF02646; RmuC; 1.
KW   DNA recombination; Coiled coil; Complete proteome.
FT   DOMAIN       16    201       COILED COIL (POTENTIAL).
SQ   SEQUENCE   453 AA;  51539 MW;  1E7EA97E82EC5E4B CRC64;

  Query Match            35.5%;  Score 49;  DB 1;  Length 453;
  Best Local Similarity  55.6%;  Pred. No. 8.6;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          3 WAAQRYGR--ELRRMSDE 18
              ||::| ||  ||||:: |
Db         65 WASERQGREEELRRLASE 82


RESULT 13
RAS3_RHIRA
ID   RAS3_RHIRA     STANDARD;      PRT;   205 AA.
AC   P22280;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-like protein 3.
GN   RAS3.
OS   Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Mucor.
OX   NCBI_TaxID=4841;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 1216B;
RX   MEDLINE=91061774; PubMed=1701021;
RA   Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
RT   "Expression of a gene family in the dimorphic fungus Mucor racemosus
RT   which exhibits striking similarity to human ras genes.";
RL   Mol. Cell. Biol. 10:6654-6663(1990).
CC   -!- ENZYME REGULATION: ALTERNATE BETWEEN AN INACTIVE FORM BOUND TO GDP
CC       AND AN ACTIVE FORM BOUND TO GTP. ACTIVATED BY A GUANINE
CC       NUCLEOTIDE-EXCHANGE FACTOR (GEF) AND INACTIVATED BY A GTPASE-
CC       ACTIVATING PROTEIN (GAP).
CC   -!- SUBCELLULAR LOCATION: Plasma membrane.
CC   -!- DEVELOPMENTAL STAGE: IN SPORULATING MYCELIUM AND MUCH LESS IN
CC       GERMLING AND YEAST.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55177; AAA83379.1; -.
DR   PIR; C36365; C36365.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      16     23       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     122    125       GTP (BY SIMILARITY).
FT   DOMAIN       38     46       EFFECTOR REGION (PROBABLE).
FT   LIPID       202    202       FARNESYL (BY SIMILARITY).
SQ   SEQUENCE   205 AA;  23408 MW;  DBF086466F090F50 CRC64;

  Query Match            34.8%;  Score 48;  DB 1;  Length 205;
  Best Local Similarity  62.5%;  Pred. No. 5.1;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         10 RELRRMSDEFEGSFKG 25
              ||:|||: | ||  ||
Db        168 REIRRMNKEQEGRSKG 183


RESULT 14
6PGL_THEMA
ID   6PGL_THEMA     STANDARD;      PRT;   220 AA.
AC   Q9X0N8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN   PGL OR DEVB OR TM1154.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC       PHOSPHOGLUCONATE.
CC   -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC       phospho-D-gluconate.
CC   -!- PATHWAY: Pentose phosphate pathway; second step.
CC   -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC       ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001772; AAD36230.1; -.
DR   PIR; F72289; F72289.
DR   TIGR; TM1154; -.
DR   InterPro; IPR006148; Gluc_gal_isom.
DR   InterPro; IPR005900; Phosphogluconlac.
DR   Pfam; PF01182; Glucosamine_iso; 1.
DR   TIGRFAMs; TIGR01198; pgl; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   220 AA;  25325 MW;  9B0FD07EE01E60C3 CRC64;

  Query Match            34.8%;  Score 48;  DB 1;  Length 220;
  Best Local Similarity  34.8%;  Pred. No. 5.5;
```

```
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          4 AAQRYGRELRRMSDEFEGSFKGL 26
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133


RESULT 15
THRC_SOLTU
ID   THRC_SOLTU     STANDARD;      PRT;   519 AA.
AC   Q9MT28;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Threonine synthase, chloroplast precursor (EC 4.2.3.1) (TS).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RX   PubMed=10940468;
RA   Casazza A.P., Basner A., Hoefgen R., Hesse H.;
RT   "Expression of threonine synthase from Solanum tuberosum L. is not
RT   metabolically regulated by photosynthesis-related signals or by
RT   nitrogenous compounds.";
RL   Plant Sci. 157:43-50(2000).
CC   -!- CATALYTIC ACTIVITY: O-phospho-L-homoserine + H(2)O = L-threonine +
CC       phosphate.
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- ENZYME REGULATION: Allosterically activated by S-adenosyl-
CC       methionine (SAM) (By similarity).
CC   -!- PATHWAY: Threonine biosynthesis; last step.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE SERINE/THREONINE DEHYDRATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF082894; AAF74984.1; -.
DR   HSSP; Q9S7B5; 1E5X.
DR   InterPro; IPR001926; B6_enzyme_beta.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   InterPro; IPR004450; Thr_synthase.
DR   Pfam; PF00291; PALP; 1.
DR   TIGRFAMs; TIGR00260; thrC; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Threonine biosynthesis; Lyase; Pyridoxal phosphate; Allosteric enzyme;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1     40       CHLOROPLAST (BY SIMILARITY).
FT   CHAIN        41    519       THREONINE SYNTHASE.
FT   BINDING     196    196       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   519 AA;  57412 MW;  114C0979CD231464 CRC64;

  Query Match            34.8%;  Score 48;  DB 1;  Length 519;
  Best Local Similarity  35.3%;  Pred. No. 14;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    8;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSD--------EFEGSFKGL 26
              ||: |:|:|::   |:|          |||| |
Db        165 NLFWAERFGKQFLGMTDLWVKHCGISHTGSFKDL 198


Search completed: September 15, 2003, 17:22:59
Job time : 7.12857 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 28.2286 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-1
Perfect score: 138
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	63.0	146	13	Q9I9N2	Q9i9n2 brachydanio
2	53	38.4	196	16	Q8VJS3	Q8vjs3 mycobacteri
3	53	38.4	223	16	Q10843	Q10843 mycobacteri
4	53	38.4	946	11	Q8K016	Q8k016 mus musculu
5	52	37.7	471	17	Q8ZY71	Q8zy71 pyrobaculum
6	51	37.0	231	10	Q9SEG0	Q9seg0 arabidopsis
7	51	37.0	232	10	Q9SQ20	Q9sq20 arabidopsis
8	51	37.0	232	10	Q9SQ22	Q9sq22 arabidopsis
9	51	37.0	232	10	Q9SQ17	Q9sq17 arabidopsis
10	51	37.0	232	10	Q9SQ19	Q9sq19 arabidopsis
11	51	37.0	232	10	Q9SQ21	Q9sq21 arabidopsis
12	51	37.0	232	10	Q8LB79	Q81b79 arabidopsis
13	51	37.0	232	10	Q9SQ16	Q9sq16 arabidopsis
14	51	37.0	232	10	Q9SQ15	Q9sq15 arabidopsis
15	51	37.0	232	10	Q9S7Q3	Q9s7q3 arabidopsis
16	51	37.0	232	10	Q9SQ18	Q9sq18 arabidopsis
17	50.5	36.6	904	2	Q9KGW3	Q9kgw3 pseudomonas
18	50.5	36.6	909	16	Q8EI34	Q8ei34 shewanella
19	50	36.2	168	11	Q8K316	Q8k316 mus musculu
20	50	36.2	283	15	O37056	O37056 chimpanzee
21	50	36.2	374	17	Q9HNZ9	Q9hnz9 halobacteri
22	50	36.2	516	10	Q9SSP5	Q9ssp5 arabidopsis
23	50	36.2	851	11	Q8K2H3	Q8k2h3 mus musculu
24	49.5	35.9	153	5	Q9UB33	Q9ub33 anopheles g
25	49.5	35.9	401	5	O97407	O97407 anopheles g
26	49.5	35.9	505	8	O47148	O47148 menziesia c
27	49.5	35.9	506	8	O47149	O47149 rhododendro
28	49.5	35.9	506	8	O47171	O47171 rhododendro
29	49.5	35.9	506	8	O63960	O63960 rhododendro
30	49.5	35.9	506	8	O62982	O62982 rhododendro
31	49.5	35.9	506	8	O62975	O62975 rhododendro
32	49.5	35.9	506	8	O62972	O62972 rhododendro
33	49.5	35.9	506	8	O62989	O62989 rhododendro
34	49.5	35.9	506	8	O62978	O62978 rhododendro
35	49.5	35.9	506	8	O47155	O47155 rhododendro
36	49.5	35.9	506	8	O47152	O47152 rhododendro
37	49.5	35.9	506	8	O47173	O47173 rhododendro
38	49.5	35.9	506	8	O62990	O62990 rhododendro
39	49.5	35.9	506	8	O62974	O62974 rhododendro
40	49.5	35.9	506	8	O62993	O62993 menziesia m
41	49.5	35.9	506	8	O47170	O47170 rhododendro
42	49.5	35.9	506	8	O47174	O47174 rhododendro
43	49.5	35.9	506	8	O62983	O62983 rhododendro
44	49.5	35.9	506	8	O62980	O62980 rhododendro
45	49.5	35.9	506	8	O62981	O62981 rhododendro

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 63.0%; Score 87; DB 13; Length 146;
Best Local Similarity 65.2%; Pred. No. 5.3e-05;
Matches 15; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 LWAAQRYGRELRRMSDEFEGSFK 24
              |||| :||::||||||||:   |
Db         89 LWAAKKYGQQLRRMSDEFDKGMK 111
```

RESULT 2
Q8VJS3
ID Q8VJS3 PRELIMINARY; PRT; 196 AA.
AC Q8VJS3;

```
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   IS1607, transposase.
GN   MT2070.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE007058; AAK46348.1; -.
DR   TIGR; MT2070; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
SQ   SEQUENCE   196 AA;  21349 MW;  C145A8D836FD9C2D CRC64;

  Query Match             38.4%;  Score 53;  DB 16;  Length 196;
  Best Local Similarity   58.8%;  Pred. No. 7.7;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        134 NLWAADRYNRAIARGHD 150


RESULT 3
Q10843
ID   Q10843      PRELIMINARY;      PRT;   223 AA.
AC   Q10843;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein Rv2014.
GN   RV2014 OR MTCY39.03C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SIMILARITY: TO M.PARATUBERCULOSIS IS900.
DR   EMBL; Z74025; CAA98415.1; -.
DR   TubercuList; Rv2014; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   223 AA;  24132 MW;  70456750017FEF37 CRC64;

  Query Match             38.4%;  Score 53;  DB 16;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 8.9;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 4
Q8K016
ID   Q8K016      PRELIMINARY;      PRT;   946 AA.
AC   Q8K016;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Inter-alpha trypsin inhibitor, heavy chain 2.
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC034341; AAH34341.1; -.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE   946 AA;  105945 MW;  8B17DBA71B85BC5C CRC64;

  Query Match             38.4%;  Score 53;  DB 11;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 44;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              |:|  :  |    : | ||| |:|:
Db        212 NVWIMEPQGMRFLHVPDTFEGHFQGV 237


RESULT 5
Q8ZY71
ID   Q8ZY71      PRELIMINARY;      PRT;   471 AA.
AC   Q8ZY71;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein PAE0922.
GN   PAE0922.
OS   Pyrobaculum aerophilum.
OC   Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC   Thermoproteaceae; Pyrobaculum.
OX   NCBI_TaxID=13773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX   PubMed=11792869;
RA   Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA   Miller J.H.;
RT   "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT   aerophilum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR   EMBL; AE009793; AAL63125.1; -.
DR   InterPro; IPR006638; Elp3.
DR   InterPro; IPR000182; GCN5acetyltransf.
DR   Pfam; PF00583; Acetyltransf; 1.
DR   SMART; SM00729; Elp3; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   471 AA;  52952 MW;  3B1E36E8AEE2EF0A CRC64;
```

```
  Query Match             37.7%;  Score 52;  DB 17;  Length 471;
  Best Local Similarity   41.7%;  Pred. No. 29;
  Matches   10;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 WAAQRYGRELRRMSDEFEGSFKGL 26
              |     |||| |:::|  | |  |
Db        404 WQHSGMGRELMRLAEEIAGEFGAL 427


RESULT 6
Q9SEG0
ID   Q9SEG0      PRELIMINARY;      PRT;   231 AA.
AC   Q9SEG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Apetala3 (Fragment).
OS   Arabidopsis lyrata.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=59689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99404148; PubMed=10474900;
RA   Lawton-Rauh A.L., Buckler E.S. IV, Purugganan M.D.;
RT   "Patterns of molecular evolution among paralogous floral homeotic
RT   genes.";
RL   Mol. Biol. Evol. 16:1037-1045(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF143380; AAF25590.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
FT   NON_TER     231    231
SQ   SEQUENCE   231 AA;  27176 MW;  A67CAE1EEBD8F7AA CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 231;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |         :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 7
Q9SQ20
ID   Q9SQ20      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ20;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Corscalla;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115806; AAD51895.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27342 MW;  BDFDCB59B73F4601 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |         :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 8
Q9SQ22
ID   Q9SQ22      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ22;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Li-8;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115801; AAD51890.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27267 MW;  42A852D697E22A65 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
```

```
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 9
Q9SQ17
ID   Q9SQ17      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ17;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115809; AAD51898.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27284 MW;  04FCFC55B73C7729 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 10
Q9SQ19
ID   Q9SQ19      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ19;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Bretagny;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115807; AAD51896.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27311 MW;  71AE593FB8A67EC3 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 11
Q9SQ21
ID   Q9SQ21      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ21;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Kent;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115805; AAD51894.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27286 MW;  66976305B88B63E3 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133
```

```
RESULT 12
Q8LB79
ID   Q8LB79      PRELIMINARY;      PRT;   232 AA.
AC   Q8LB79;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral homeotic protein APETALA3 (AP3).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-Length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AY087369; AAM64919.1; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27339 MW;  CC90703F959CFAD5 CRC64;

  Query Match            37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLNKLDIQELRRLEDEMENTFK 133


RESULT 13
Q9SQ16
ID   Q9SQ16      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ16;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115810; AAD51899.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27314 MW;  DB8CA1FC835557D6 CRC64;

  Query Match            37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 14
Q9SQ15
ID   Q9SQ15      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ15;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Kas-1;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115812; AAD51901.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27300 MW;  5CA05FD44F824DF0 CRC64;

  Query Match            37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRVGECLDELDIQELRRLEDEMENTFK 133
```

```
RESULT 15
Q9S7Q3
ID   Q9S7Q3      PRELIMINARY;      PRT;   232 AA.
AC   Q9S7Q3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. GR-3, and cv. CHI-1;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115803; AAD51892.1; -.
DR   EMBL; AF115798; AAD51887.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27340 MW;  6690703F9F9CFD63 CRC64;

  Query Match             37.0%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 18;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRLGECLDKLDIQELRRLEDEMENTFK 133


Search completed: September 15, 2003, 17:25:45
Job time : 30.2286 secs
```

V for consen.? No

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2003, 09:43:16 ; Search time 86 Seconds
(without alignments)
49.833 Million cell updates/sec

Title: US-09-544-664B-2
Perfect score: 142
Sequence: 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	137	96.5	166	18	AAW32476	BBC6 protein for r
2	137	96.5	168	19	AAW55779	Human Bcl-xL/Bcl-2
3	137	96.5	168	21	AAB13512	Human cell prolife
4	137	96.5	168	22	AAB70368	Human BAD mutant a
5	137	96.5	168	22	AAB48287	Human Bad protein.
6	137	96.5	168	22	AAG67688	Amino acid sequenc
7	137	96.5	168	24	ABR39081	Human BAD protein
8	137	96.5	201	23	ABP41630	Human ovarian anti
9	127	89.4	25	23	ABP56161	PTPC-interacting T
10	127	89.4	25	23	ABG78484	Mutant Bcl2 compet
11	127	89.4	25	23	AAU78610	Human Bad peptide
12	124	87.3	25	23	ABG78490	Mutant Bcl2 compet
13	124	87.3	25	23	AAU78617	Human Bad peptide
14	123	86.6	25	23	ABG78488	Mutant Bcl2 compet
15	123	86.6	25	23	ABG78489	Mutant Bcl2 compet
16	123	86.6	25	23	AAU78615	Human Bad peptide
17	123	86.6	25	23	AAU78616	Human Bad peptide
18	122	85.9	24	23	AAU78627	Human Bad peptide
19	122	85.9	25	23	ABG78486	Mutant Bcl2 compet
20	122	85.9	25	23	ABG78493	Mutant Bcl2 compet
21	122	85.9	25	23	ABG78497	Mutant Bcl2 compet
22	122	85.9	25	23	AAU78612	Human Bad peptide
23	122	85.9	25	23	AAU78620	Human Bad peptide
24	122	85.9	25	23	AAU78624	Human Bad peptide
25	121	85.2	25	23	ABG78485	Mutant Bcl2 compet
26	121	85.2	25	23	ABG78492	Mutant Bcl2 compet
27	121	85.2	25	23	AAU78611	Human Bad peptide
28	121	85.2	25	23	AAU78619	Human Bad peptide
29	119	83.8	25	23	ABG78491	Mutant Bcl2 compet
30	119	83.8	25	23	ABG78498	Mutant Bcl2 compet
31	119	83.8	25	23	AAU78618	Human Bad peptide
32	119	83.8	25	23	AAU78625	Human Bad peptide
33	117	82.4	23	23	AAU78628	Human Bad peptide
34	116	81.7	26	21	AAY96321	Mammalian Bad Bcl-
35	116	81.7	26	22	AAB70371	BAD BH3 consensus
36	113	79.6	25	23	ABG78487	Mutant Bcl2 compet
37	113	79.6	25	23	AAU78613	Human Bad peptide
38	112	78.9	25	23	ABG78495	Mutant Bcl2 compet
39	112	78.9	25	23	AAU78622	Human Bad peptide
40	111	78.2	22	23	AAU78629	Human Bad peptide
41	110	77.5	25	23	ABG78496	Mutant Bcl2 compet
42	110	77.5	25	23	AAU78623	Human Bad peptide
43	109	76.8	26	21	AAB37001	Bcl2 polypeptide B
44	109	76.8	26	21	AAB37002	Bcl2 polypeptide B
45	109	76.8	27	21	AAB37003	Bcl2 polypeptide B

ALIGNMENTS

RESULT 1
AAW32476
ID AAW32476 standard; Protein; 166 AA.
XX
AC AAW32476;
XX
DT 15-JAN-1998 (first entry)
XX
DE BBC6 protein for regulating cell death.
XX
KW BBC6 gene; cell death; cell cycle; Bcl2; human.
XX
OS Homo sapiens.
XX
PN US5663316-A.
XX
PD 02-SEP-1997.
XX
PF 18-JUN-1996; 96US-0665617.
XX
PR 18-JUN-1996; 96US-0665617.
XX
PA (CLON-) CLONTECH LAB INC.
XX
PI Xudong Y;
XX
DR WPI; 1997-447980/41.
DR N-PSDB; AAT91561.
XX
PT Isolated BBC6 gene - encodes a protein that regulates cell death
PT through interaction with Bcl-2
XX

PS Claim 1; Column 11-12; 7pp; English.
XX
CC The present sequence represents a protein of 166 amino acids. The
CC sequence is disclosed as being a protein called BBC6 which regulates
CC cell death through interaction with Bcl-2. The DNA may be used for the
CC production of the recombinant protein, which can be used in unspecified
CC therapeutic or diagnostic procedures, as a molecular weight marker, and
CC to raise antibodies that can be used in unspecified diagnostic or
CC therapeutic applications and to reduce or eliminate the biological
CC activity of the BBC6 protein in vivo.
XX
SQ Sequence 166 AA;

Query Match 96.5%; Score 137; DB 18; Length 166;
Best Local Similarity 96.3%; Pred. No. 1.1e-12;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
Db 101 NLWAAQRYGRELRRMSDEFVDSFKKGL 127

RESULT 2
AAW55779
ID AAW55779 standard; Protein; 168 AA.
XX
AC AAW55779;
XX
DT 17-JUL-1998 (first entry)
XX
DE Human Bcl-xL/Bcl-2 associated death promoting polypeptide.
XX
KW Human; Bcl-xL/Bcl-2 associated death promoting polypeptide; Bad;
KW programmed cell death; apoptosis.
XX
OS Homo sapiens.
XX
PN WO9812328-A2.
XX
PD 26-MAR-1998.
XX
PF 18-SEP-1997; 97WO-US16991.
XX
PR 20-SEP-1996; 96US-0717123.
XX
PA (IDUN-) IDUN PHARM INC.
XX
PI Horne WA, Oltersdorf T;
XX
DR WPI; 1998-217267/19.
DR N-PSDB; AAV25877.
XX
PT Bad gene mediating apoptosis - used to develop products for treating
PT e.g. neurodegenerative disease, cancers or autoimmune disease
XX
PS Claim 8; Fig 1; 41pp; English.
XX
CC The present sequence is the human Bcl-xL/Bcl-2 associated
CC death promoting polypeptide, Bad, the binding of which to Bcl-Xl
CC results in the induction of programmed cell death, i.e. apoptosis.
CC Bad can be used in screening assays for compounds to treat or
CC prevent diseases characterised by apoptotic cell death, such as
CC neurodegenerative disorders, e.g. Alzheimer's and Parkinson's
CC disease, amyotrophic lateral sclerosis, retinitis pigmentosa and
CC cerebellar degeneration, and myelodysplastic syndromes, e.g.
CC aplastic anaemia and ischaemic injury including myocardial
CC infarction, stroke and reperfusion injury. Assays can also be
CC used to obtain apoptosis enhancing compounds to treat or prevent
CC diseases characterised by the loss of apoptotic cell death, such as
CC cancers, e.g. lymphoma and hormone dependent tumours, autoimmune
CC diseases, e.g. systemic lupus erythematosus and immune-mediated
CC glomerulonephritis and viral infections, e.g. herpesvirus,
CC poxvirus or adenovirus infection. Bad can also be used for
CC detection and diagnosis.
XX
SQ Sequence 168 AA;

Query Match 96.5%; Score 137; DB 19; Length 168;
Best Local Similarity 96.3%; Pred. No. 1.1e-12;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
Db 103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129

RESULT 3
AAB13512
ID AAB13512 standard; protein; 168 AA.
XX
AC AAB13512;
XX
DT 02-NOV-2000 (first entry)
XX
DE Human cell proliferation protein APOP-1.
XX
KW Human; cell proliferation; APOP-1; cancer; inflammation; infection;
KW trauma; neurodegenerative disease; ischaemic injury; wasting disease.
XX
OS Homo sapiens.
XX
PN US6080847-A.
XX
PD 27-JUN-2000.
XX
PF 04-DEC-1997; 97US-0985335.
XX
PR 04-DEC-1997; 97US-0985335.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Corley NC, Hillman JL, Yue H, Lal P, Shah P;
XX
DR WPI; 2000-451230/39.
DR N-PSDB; AAA63332.
XX
PT Novel polynucleotide and polypeptide sequences of proteins associated
PT with cell proliferation for diagnosis, prevention and treatment of e.g.
PT cancer, acquired immunodeficiency syndrome, and Parkinson's disease -
XX
PS Example 8; Fig 1; 58pp; English.
XX
CC The present sequence is the human APOP-1 protein. This protein, which
CC shares structural and chemical homology with Bcl-2, is involved in cell
CC proliferation. Its coding sequence was isolated by screening a synovial
CC tissue cDNA library using a computer search for amino acid sequence
CC alignments. The gene and protein can be used in the treatment of various
CC cancers, disorders with associated inflammation such as Addison's
CC disease, adult respiratory distress syndrome, allergies, anaemia, asthma,
CC atherosclerosis, Crohn's disease, ulcerative colitis, diabetes mellitus,
CC emphysema, glomerulonephritis, gout, Graves' disease, irritable bowel
CC syndrome, lupus erythematosus, multiple sclerosis, myasthenia gravis,
CC myocardial or pericardial inflammation, osteoporosis, rheumatoid
CC arthritis, Sjogren's syndrome and autoimmune thyroiditis, complications
CC of cancer, haemodialysis and extracorporeal circulation, infections,
CC trauma, disorders with associated apoptosis including AIDS and other
CC infectious and genetic immunodeficiencies, neurodegenerative diseases
CC such as Alzheimer's disease and Parkinson's disease, ischaemic injuries
CC such as myocardial infarction, and wasting diseases including cachexia.
XX
SQ Sequence 168 AA;

Query Match 96.5%; Score 137; DB 21; Length 168;
Best Local Similarity 96.3%; Pred. No. 1.1e-12;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129
```

RESULT 4
AAB70368
ID AAB70368 standard; protein; 168 AA.
XX
AC AAB70368;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human BAD mutant amino acid sequence SEQ ID NO:1.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 1; Page 147; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed human BAD mutant amino acid sequence from the present invention.
XX
SQ Sequence 168 AA;

```
  Query Match             96.5%;  Score 137;  DB 22;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.1e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129
```

RESULT 5
AAB48287
ID AAB48287 standard; protein; 168 AA.
XX
AC AAB48287;
XX
DT 02-APR-2001 (first entry)
XX
DE Human Bad protein.
XX
KW S-phase kinase associated protein; SKP1; SKP2; SKP2-like protein; ZF;
KW CUL-1; cullin; CDC53; p27; cyclin E; Max; Mad; c-Myc; MDM2; p53; Bax;
KW Bad; Bcl-2; tumour; cytostatic.
XX
OS Homo sapiens.
XX
PN WO200075184-A1.
XX
PD 14-DEC-2000.
XX
PF 05-JUN-2000; 2000WO-US15449.
XX
PR 04-JUN-1999; 99US-0137494.
XX
PA (UYYA) UNIV YALE.
XX
PI Zhang H, Tsvetkov LM, Kondo T;
XX
DR WPI; 2001-061703/07.
DR N-PSDB; AAC84599.
XX
PT Modulating polypeptide levels in a cell, diagnosing and treating tumor,
PT involves altering levels of proteins such as S-phase kinase associated
PT proteins 1, 2 and cullin/CDC53 proteins -
XX
PS Claim 5; Page 102-103; 162pp; English.
XX
CC The invention relates to methods of altering the polypeptide levels in a
CC cell, using proteins selected from S-phase kinase associated proteins 1
CC and 2 (SKP1, SKP2), SKP2-like proteins (ZF) and CUL-1 (a member of the
CC cullin/ CDC53 family of proteins). The method is useful for altering the
CC level of p27, cyclin E, Max, Mad, c-Myc, MDM2, p53, Bax, Bad or Bcl-2
CC polypeptide in a cell. SKP2 and SKP2-like protein levels are useful for
CC detecting tumours, and in monitoring tumor treatment in a mammal. Agents
CC that modulate interactions between SKP and target proteins are useful for
CC treating tumours.
XX
SQ Sequence 168 AA;

```
  Query Match             96.5%;  Score 137;  DB 22;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.1e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129
```

RESULT 6
AAG67688
ID AAG67688 standard; Protein; 168 AA.
XX
AC AAG67688;
XX
DT 26-NOV-2001 (first entry)
XX
DE Amino acid sequence of protein associated with cell proliferation-1.
XX
KW Human; cell proliferation; APOP-1; APOP-2; APOP-3; apoptosis; cancer;

```
KW   brain cancer; breast cancer; Alzheimer's disease; Parkinson's disease;
KW   inflammation; allergy; gout; osteoarthritis; bronchitis.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   10..13
FT                   /note= "potential casein kinase II phosphorylation site"
FT   Modified-site   16..19
FT                   /note= "potential casein kinase II phosphorylation site"
FT   Modified-site   34..36
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Modified-site   80..83
FT                   /note= "potential casein kinase II phosphorylation site"
FT   Modified-site   115..118
FT                   /note= "potential cAMP- and cGMP-dependent protein
FT                   kinase phosphorylation site"
FT   Modified-site   124..126
FT                   /note= "potential protein kinase C phosphorylation site"
FT   Modified-site   153..156
FT                   /note= "potential casein kinase II phosphorylation site"
XX
PN   US6281334-B1.
XX
PD   28-AUG-2001.
XX
PF   30-SEP-1999;   99US-0410372.
XX
PR   04-DEC-1997;   97US-0985335.
XX
PA   (INCY-) INCYTE GENOMICS INC.
XX
PI   Hillman JL,  Yue H,  Lal P,  Shah P,  Corley NC;
XX
DR   WPI; 2001-569961/64.
DR   N-PSDB; AAH78430.
XX
PT   New polypeptides associated with cell proliferation, useful for
PT   preventing or treating cancer (e.g. brain cancer), a disorder
PT   associated with an increase in apoptosis (e.g. Alzheimer's disease) or
PT   inflammation (e.g. gout)  -
XX
PS   Example; Fig 1A-C; 59pp; English.
XX
CC   The present sequence represents a human protein which is associated
CC   with cell proliferation, designated APOP-1. The specification also
CC   describes APOP-2 and APOP-3. The APOP polypeptides are useful for
CC   diagnosing, preventing or treating disorders associated with abnormal
CC   cell proliferation and apoptosis. The polypeptides and composition are
CC   particularly useful for treating or preventing cancer (e.g. brain or
CC   breast cancer), a disorder associated with an increase in apoptosis
CC   (e.g. Alzheimer's disease or Parkinson's disease) or inflammation
CC   (e.g. allergies, gout, osteoarthritis or bronchitis).
XX
SQ   Sequence   168 AA;

  Query Match            96.5%;  Score 137;  DB 22;  Length 168;
  Best Local Similarity  96.3%;  Pred. No. 1.1e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 7
ABR39081
ID   ABR39081 standard; Protein; 168 AA.
XX
AC   ABR39081;
XX
DT   10-MAY-2003  (first entry)
XX
DE   Human BAD protein SEQ ID NO:2.
XX
KW   Human; BAD; herpes simplex virus; HSV; US3; herpes virus; apoptosis;
KW   virucide; infection.
XX
OS   Homo sapiens.
XX
PN   WO2003012049-A2.
XX
PD   13-FEB-2003.
XX
PF   31-JUL-2002; 2002WO-US24177.
XX
PR   31-JUL-2001; 2001US-308929P.
XX
PA   (UYCH-) UNIV CHICAGO.
XX
PI   Munger J,  Roizman B;
XX
DR   WPI; 2003-248168/24.
DR   N-PSDB; ABZ81200.
XX
PT   Inducing apoptosis in a cell infected with herpes simplex virus, HSV,
PT   by administering to the cell, a composition comprising an agent that
PT   inhibits phosphorylation of pro-apoptotic polypeptide BAD by HSV US3  -
XX
PS   Claim 15; Page 166-167; 192pp; English.
XX
CC   The present invention describes a method (M1) for inducing apoptosis in
CC   a cell infected with herpes simplex virus (HSV), which comprises
CC   administering to the cell, a composition having an agent that inhibits
CC   phosphorylation of pro-apoptotic polypeptide BAD by HSV US3. Also
CC   described is a method (M2) for treating a patient infected with HSV, by
CC   administering to the patient, a composition comprising a peptide
CC   comprising a sequence of 4-100 continuous amino acids of a 168 residue
CC   amino acid sequence (see ABR39081), where the peptide comprises ser112,
CC   ser135, or ser155, or their combinations. BAD has virucide activity.
CC   M1 is useful for inducing apoptosis in a cell infected with HSV, where
CC   the cell is in a human. M2 is useful for treating a patient infected
CC   with HSV. The present sequence represents human BAD, which is used in
CC   the exemplification of the present invention.
XX
SQ   Sequence   168 AA;

  Query Match            96.5%;  Score 137;  DB 24;  Length 168;
  Best Local Similarity  96.3%;  Pred. No. 1.1e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 8
ABP41630
ID   ABP41630 standard; Protein; 201 AA.
XX
AC   ABP41630;
XX
DT   22-AUG-2002  (first entry)
XX
DE   Human ovarian antigen HCE4K28, SEQ ID NO:2762.
XX
KW   Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
KW   ovarian cancer; breast cancer; tumour; reproductive system disorder;
KW   infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
KW   PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;
KW   inflammatory condition; immune disorder; blood disorder;
KW   cardiovascular disorder; respiratory disorder; neurological disorder;
KW   gastrointestinal disorder; urinary system disorder; drug screening;
KW   gene therapy; chromosome mapping; forensic analysis;
KW   antibody preparation; cytostatic; immunomodulatory; neuroprotective;
KW   antiinflammatory; gynaecological; reproductive.
```

XX
OS Homo sapiens.
XX
PN WO200200677-A1.
XX
PD 03-JAN-2002.
XX
PF 07-JUN-2001; 2001WO-US18569.
XX
PR 07-JUN-2000; 2000US-209467P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Birse CE, Rosen CA;
XX
DR WPI; 2002-147878/19.
DR N-PSDB; ABQ54707.
XX
PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
PT useful in the prevention, treatment and diagnosis of cancer (e.g.
PT ovarian cancer), immune disorders, cardiovascular disorders and
PT neurological diseases -
XX
PS Claim 11; SEQ ID No 2762; 2922pp; English.
XX
CC The invention relates to 2175 novel human ovarian antigens (ABP41054-
CC ABP43228) and to cDNAs encoding them (ABQ54131-ABQ56305), and also
CC encompasses polypeptides 90% identical and polynucleotides 95% identical
CC to the sequences of the invention. The invention additionally relates to
CC recombinant vectors and host cells comprising human ovarian antigen
CC polynucleotides, antibodies against human ovarian antigens, and the use
CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
CC treating, prognosing or preventing various ovary and/or breast-related
CC disorders. Such conditions include ovarian cancer and breast cancer, and
CC metastatic tumours of ovarian or breast origin, reproductive system
CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
CC vaginitis), immune disorders (e.g., congenital and acquired
CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
CC respiratory disorders, neurological disorders, gastrointestinal disorders
CC and urinary system disorders. Ovarian antigen polypeptides and
CC polynucleotides may also be used in screening for compounds which
CC modulate ovarian antigen expression or activity. The polynucleotides may
CC further be used for gene therapy, chromosome mapping, in the
CC identification of individuals and in forensic analysis, and the
CC polypeptides may be used as food additives or to prepare antibodies
CC useful in disease diagnosis, drug targeting and phenotyping. The present
CC sequence represents a human ovarian antigen of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 201 AA;

Query Match 96.5%; Score 137; DB 23; Length 201;
Best Local Similarity 96.3%; Pred. No. 1.3e-12;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
||||||| ||||||||||||||||||||
Db 111 NLWAAQRYGRELRRMSDEFVDSFKKGL 137

RESULT 9
ABP56161
ID ABP56161 standard; peptide; 25 AA.
XX
AC ABP56161;
XX
DT 28-MAR-2003 (first entry)
XX
DE PTPC-interacting TOX peptide #27.
XX
KW Mitochondrial membrane permeabilisation; mitochondrion; PTPC;
KW permeability transition pore complex; virucide; neuroprotective;
KW vasotropic; cytostatic; infection; cell death regulation; apoptosis;
KW mitochondrial permeability transition pore complex modulator; cancer;
KW apoptogenic; ischaemia; neurodegenerative disease; fulminant hepatitis.
XX
OS Synthetic.
XX
PN WO200261105-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-EP01633.
XX
PR 02-FEB-2001; 2001US-265594P.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CENT NAT RECH SCI.
XX
PI Edelman L, Jacotot E, Briand J;
XX
DR WPI; 2002-619260/66.
XX
PT New chimeric bifunctional molecules that target specific cells and
PT regulate the apoptosis function of the permeability transition pore
PT complex of the mitochondria, useful for treating or preventing e.g.
PT cancer or ischemia -
XX
PS Claim 9; Page 11; 76pp; English.
XX
CC The present invention describes a chimeric bifunctional molecule (I)
CC comprising at least a first functional molecule covalently linked to a
CC second functional molecule, which is able to modulate the activity of
CC the permeability transition pore complex (PTPC) of the mitochondria.
CC (I) has the function of specifically targeting and entering a tissue
CC cell population. The second functional molecule has the function of
CC specifically targeting, and inducing or preventing the death of the
CC cells by apoptosis by regulating the opening or the closing of the PTPC
CC of the mitochondria or its fragment. (I) has virucide, neuroprotective,
CC vasotropic and cytostatic activities, and can be used as a mitochondrial
CC permeability transition pore complex (PTPC) modulator. (I) is useful for
CC treating or preventing a pathological infection or disease. (I) is also
CC useful for regulating cell death regulatory molecules, specifically the
CC apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia,
CC neurodegenerative diseases, fulminant hepatitis or viral infections.
CC The present sequence represents a PTPC-interacting TOX peptide which is
CC given in the exemplification of the present invention.
XX
SQ Sequence 25 AA;

Query Match 89.4%; Score 127; DB 23; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.2e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKK 25
||||| ||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFVDSFKK 25

RESULT 10
ABG78484
ID ABG78484 standard; Peptide; 25 AA.
XX
AC ABG78484;
XX
DT 15-NOV-2002 (first entry)
XX
DE Mutant Bcl2 competitive binding assay peptide #1.
XX
KW Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.

XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200240530-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US45693.
XX
PR 20-NOV-2000; 2000US-0716395.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Petros AM, Yoon H, Nettesheim DG;
XX
DR WPI; 2002-490141/52.
XX
PT New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT useful in biological assays to identify substances that block the
PT ability of Bcl-2 to inhibit programmed cell death or apoptosis -
XX
PS Example 2; Page 17; 36pp; English.
XX
CC This invention relates to a novel mutant protein which is derived from
CC a wild type human Bcl-2 protein. The mutant is created by replacing a
CC sequence of amino acid residues comprising a flexible loop from the wild
CC type Bcl-2 protein with an amino acid sequence comprising at least two
CC acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC shown in the specification. The invention also comprises an assay for
CC identifying substances that bind to the Bcl-2 protein. The protein
CC sequences of the invention are useful in biological assays to identify
CC substances that block the ability of Bcl-2 to inhibit programmed cell
CC death or apoptosis. The present sequence represents a human Bcl2
CC peptide of the invention.
XX
SQ Sequence 25 AA;

Query Match 89.4%; Score 127; DB 23; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.2e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKK 25
|||||| ||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFVDSFKK 25

RESULT 11
AAU78610
ID AAU78610 standard; Peptide; 25 AA.
XX
AC AAU78610;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Bad peptide #10 which binds to a member of the Bcl-2 family.
XX
KW Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW ischemic injury; suppressor; BH3 domain.
XX
OS Homo sapiens.
XX
PN WO200220568-A2.
XX
PD 14-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US27410.
XX
PR 06-SEP-2000; 2000US-0656399.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Joseph MK, Olejniczak ET, Petros AM;
PI Nettesheim DG, Swift KM, Matayoshi E, Zhang H;
XX
DR WPI; 2002-292254/33.
XX
PT New derivatives of Bad peptide, useful for identifying compounds that
PT bind to Bcl-2 proteins, potential agents for treating cancer and
PT degenerative disease -
XX
PS Claim 18; Page 18; 31pp; English.
XX
CC The present invention relates to new peptides that are derived from a
CC wild-type human Bad peptide and are able to bind to a member of the
CC Bcl-2 protein family. The peptides are useful, when labelled, in
CC competitive/displacement assays for identifying substances that bind to
CC members of the Bcl-2 family and may induce or suppress apoptosis so are
CC potentially useful for treating cancer (inducers) or degenerative
CC diseases or ischemic injury (suppressors). The peptides of the invention
CC have high helix propensity, maintain the contacts of the wild-type Bad
CC peptide and, compared with the Bad peptide, may have better physical
CC properties, particularly solubility. The present sequence represents one
CC of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC from the BH3 domain of the human wild-type Bad peptide.
XX
SQ Sequence 25 AA;

Query Match 89.4%; Score 127; DB 23; Length 25;
Best Local Similarity 96.0%; Pred. No. 4.2e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKK 25
|||||| ||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFVDSFKK 25

RESULT 12
ABG78490
ID ABG78490 standard; Peptide; 25 AA.
XX
AC ABG78490;
XX
DT 15-NOV-2002 (first entry)
XX
DE Mutant Bcl2 competitive binding assay peptide #7.
XX
KW Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200240530-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US45693.
XX
PR 20-NOV-2000; 2000US-0716395.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Petros AM, Yoon H, Nettesheim DG;
XX
DR WPI; 2002-490141/52.
XX
PT New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT useful in biological assays to identify substances that block the
PT ability of Bcl-2 to inhibit programmed cell death or apoptosis -
XX
PS Example 2; Page 17; 36pp; English.
XX
CC This invention relates to a novel mutant protein which is derived from
CC a wild type human Bcl-2 protein. The mutant is created by replacing a
CC sequence of amino acid residues comprising a flexible loop from the wild
CC type Bcl-2 protein with an amino acid sequence comprising at least two

```
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   25 AA;

  Query Match            87.3%;  Score 124;  DB 23;  Length 25;
  Best Local Similarity  92.0%;  Pred. No. 1.2e-11;
  Matches   23;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKK 25
              |||||| |||||||||||||||:|||
Db          1 NLWAAQRYGRELRRMSDEFVDAFKK 25


RESULT 13
AAU78617
ID   AAU78617 standard; Peptide; 25 AA.
XX
AC   AAU78617;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human Bad peptide #17 which binds to a member of the Bcl-2 family.
XX
KW   Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW   ischemic injury; suppressor; BH3 domain.
XX
OS   Homo sapiens.
XX
PN   WO200220568-A2.
XX
PD   14-MAR-2002.
XX
PF   04-SEP-2001; 2001WO-US27410.
XX
PR   06-SEP-2000; 2000US-0656399.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Joseph MK,  Olejniczak ET,  Petros AM;
PI   Nettesheim DG,  Swift KM,  Matayoshi E,  Zhang H;
XX
DR   WPI; 2002-292254/33.
XX
PT   New derivatives of Bad peptide, useful for identifying compounds that
PT   bind to Bcl-2 proteins, potential agents for treating cancer and
PT   degenerative disease  -
XX
PS   Example 1; Page 14; 31pp; English.
XX
CC   The present invention relates to new peptides that are derived from a
CC   wild-type human Bad peptide and are able to bind to a member of the
CC   Bcl-2 protein family. The peptides are useful, when labelled, in
CC   competitive/displacement assays for identifying substances that bind to
CC   members of the Bcl-2 family and may induce or suppress apoptosis so are
CC   potentially useful for treating cancer (inducers) or degenerative
CC   diseases or ischemic injury (suppressors). The peptides of the invention
CC   have high helix propensity, maintain the contacts of the wild-type Bad
CC   peptide and, compared with the Bad peptide, may have better physical
CC   properties, particularly solubility. The present sequence represents one
CC   of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC   from the BH3 domain of the human wild-type Bad peptide.
XX
SQ   Sequence   25 AA;

  Query Match            87.3%;  Score 124;  DB 23;  Length 25;
  Best Local Similarity  92.0%;  Pred. No. 1.2e-11;
  Matches   23;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKK 25
              |||||| |||||||||||||||:|||
Db          1 NLWAAQRYGRELRRMSDEFVDAFKK 25


RESULT 14
ABG78488
ID   ABG78488 standard; Peptide; 25 AA.
XX
AC   ABG78488;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mutant Bcl2 competitive binding assay peptide #5.
XX
KW   Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200240530-A2.
XX
PD   23-MAY-2002.
XX
PF   15-NOV-2001; 2001WO-US45693.
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 17; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   25 AA;

  Query Match            86.6%;  Score 123;  DB 23;  Length 25;
  Best Local Similarity  92.0%;  Pred. No. 1.7e-11;
  Matches   23;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKK 25
              ||| || ||||||||||||||||||||
Db          1 NLWGAQRYGRELRRMSDEFVDSFKK 25


RESULT 15
ABG78489
ID   ABG78489 standard; Peptide; 25 AA.
XX
AC   ABG78489;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mutant Bcl2 competitive binding assay peptide #6.
```

XX
KW Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200240530-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US45693.
XX
PR 20-NOV-2000; 2000US-0716395.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Petros AM, Yoon H, Nettesheim DG;
XX
DR WPI; 2002-490141/52.
XX
PT New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT useful in biological assays to identify substances that block the
PT ability of Bcl-2 to inhibit programmed cell death or apoptosis -
XX
PS Example 2; Page 17; 36pp; English.
XX
CC This invention relates to a novel mutant protein which is derived from
CC a wild type human Bcl-2 protein. The mutant is created by replacing a
CC sequence of amino acid residues comprising a flexible loop from the wild
CC type Bcl-2 protein with an amino acid sequence comprising at least two
CC acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC shown in the specification. The invention also comprises an assay for
CC identifying substances that bind to the Bcl-2 protein. The protein
CC sequences of the invention are useful in biological assays to identify
CC substances that block the ability of Bcl-2 to inhibit programmed cell
CC death or apoptosis. The present sequence represents a human Bcl2
CC peptide of the invention.
XX
SQ Sequence 25 AA;

Query Match 86.6%; Score 123; DB 23; Length 25;
Best Local Similarity 92.0%; Pred. No. 1.7e-11;
Matches 23; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy          1 NLWAAQEYGRELRRMSDEFVDSFKK 25
              |||| | ||||||||||||||||||||
Db          1 NLWAGQRYGRELRRMSDEFVDSFKK 25
```

Search completed: September 23, 2003, 09:47:13
Job time : 87 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 23, 2003, 09:43:16 ; Search time 30 Seconds
(without alignments)
38.080 Million cell updates/sec

Title: US-09-544-664B-2
Perfect score: 142
Sequence: 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	137	96.5	166	1	US-08-665-617-2	Sequence 2, Appli
2	137	96.5	168	2	US-08-717-123-2	Sequence 2, Appli
3	137	96.5	168	3	US-08-985-335-1	Sequence 1, Appli
4	137	96.5	168	3	US-08-985-335-7	Sequence 7, Appli
5	137	96.5	168	3	US-09-410-372-1	Sequence 1, Appli
6	137	96.5	168	3	US-09-410-372-7	Sequence 7, Appli
7	137	96.5	168	4	US-09-375-257-2	Sequence 2, Appli
8	109	76.8	204	1	US-08-333-565-2	Sequence 2, Appli
9	109	76.8	204	2	US-08-661-479-2	Sequence 2, Appli
10	109	76.8	204	2	US-08-733-505A-1	Sequence 1, Appli
11	109	76.8	204	2	US-08-733-505A-12	Sequence 12, Appl
12	109	76.8	204	2	US-08-733-505A-13	Sequence 13, Appl
13	109	76.8	204	2	US-08-733-505A-14	Sequence 14, Appl
14	106	74.6	204	2	US-08-717-123-3	Sequence 3, Appli
15	106	74.6	204	4	US-09-375-257-3	Sequence 3, Appli
16	97	68.3	23	1	US-08-333-565-10	Sequence 10, Appl
17	97	68.3	23	2	US-08-661-479-10	Sequence 10, Appl
18	97	68.3	59	2	US-08-733-505A-55	Sequence 55, Appl
19	97	68.3	59	2	US-08-733-505A-56	Sequence 56, Appl
20	97	68.3	59	2	US-08-733-505A-57	Sequence 57, Appl
21	97	68.3	59	2	US-08-733-505A-58	Sequence 58, Appl
22	81	57.0	16	1	US-08-333-565-26	Sequence 26, Appl
23	81	57.0	16	2	US-08-661-479-26	Sequence 26, Appl
24	56	39.4	11	2	US-08-733-505A-34	Sequence 34, Appl
25	56	39.4	11	2	US-08-706-741B-69	Sequence 69, Appl
26	56	39.4	11	2	US-08-924-695A-69	Sequence 69, Appl
27	51	35.9	1125	4	US-09-252-991A-18729	Sequence 18729, A
28	48	33.8	626	4	US-09-252-991A-26276	Sequence 26276, A
29	48	33.8	1064	4	US-09-252-991A-17508	Sequence 17508, A
30	47	33.1	467	4	US-09-252-991A-18296	Sequence 18296, A
31	46	32.4	347	2	US-08-379-556A-2	Sequence 2, Appli
32	44	31.0	66	2	US-08-867-087B-40	Sequence 40, Appl
33	44	31.0	124	3	US-08-591-632-56	Sequence 56, Appl
34	44	31.0	124	4	US-09-611-451-56	Sequence 56, Appl
35	44	31.0	371	4	US-09-252-991A-20008	Sequence 20008, A
36	44	31.0	587	3	US-09-147-923-2	Sequence 2, Appli
37	44	31.0	724	4	US-09-328-352-7710	Sequence 7710, Ap
38	43.5	30.6	529	4	US-09-252-991A-22964	Sequence 22964, A
39	43.5	30.6	1047	4	US-09-198-452A-1058	Sequence 1058, Ap
40	43	30.3	125	4	US-09-328-352-7449	Sequence 7449, Ap
41	43	30.3	284	4	US-09-328-352-6559	Sequence 6559, Ap
42	43	30.3	2008	4	US-09-091-501B-8	Sequence 8, Appli
43	43	30.3	3433	4	US-09-091-501B-10	Sequence 10, Appl
44	42.5	29.9	460	4	US-09-252-991A-23461	Sequence 23461, A
45	42	29.6	28	1	US-08-261-660A-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
; GENERAL INFORMATION:
; APPLICANT: Xudong, Yin
; TITLE OF INVENTION: Gene and Protein for Regulation of Cell Death
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: Florida
; COUNTRY: USA
; ZIP: 32606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,617
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Saliwanchik, David R.
; REGISTRATION NUMBER: 31,794
; REFERENCE/DOCKET NUMBER: CL-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (352) 375-8100
; TELEFAX: (352) 372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 166 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-665-617-2

Query Match 96.5%; Score 137; DB 1; Length 166;
Best Local Similarity 96.3%; Pred. No. 5.3e-14;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        101 NLWAAQRYGRELRRMSDEFVDSFKKGL 127
```

RESULT 2

```
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match             96.5%;  Score 137;  DB 2;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 3
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match             96.5%;  Score 137;  DB 3;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 4
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match             96.5%;  Score 137;  DB 3;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 5
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match             96.5%;  Score 137;  DB 3;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 6
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match             96.5%;  Score 137;  DB 3;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 7
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
```

```
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match             96.5%;  Score 137;  DB 4;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 5.4e-14;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 8
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-333-565-2

  Query Match             76.8%;  Score 109;  DB 1;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| ||||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 9
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             76.8%;  Score 109;  DB 2;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| ||||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 10
US-08-733-505A-1
; Sequence 1, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
```

```
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-733-505A-1

  Query Match             76.8%;  Score 109;  DB 2;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 11
US-08-733-505A-12
; Sequence 12, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-12

  Query Match             76.8%;  Score 109;  DB 2;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 12
US-08-733-505A-13
; Sequence 13, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-13

  Query Match             76.8%;  Score 109;  DB 2;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 13
US-08-733-505A-14
; Sequence 14, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
```

```
;       STREET:  7733 FORSYTH BLVD., SUITE 1400
;       CITY:  ST. LOUIS
;       STATE:  MISSOURI
;       COUNTRY:  USA
;       ZIP:  63105
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.30
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/733,505A
;       FILING DATE:
;       CLASSIFICATION:  530
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  HOLLAND, DONALD R.
;       REGISTRATION NUMBER:  35,197
;       REFERENCE/DOCKET NUMBER:  965458
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (314) 727-5188
;       TELEFAX:  (314) 727-6092
;   INFORMATION FOR SEQ ID NO:  14:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  204 amino acids
;       TYPE:  amino acid
;       STRANDEDNESS:
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  peptide
US-08-733-505A-14

  Query Match             76.8%;  Score 109;  DB 2;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.7e-09;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 14
US-08-717-123-3
; Sequence 3, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;     APPLICANT:  Horne, William A.
;     APPLICANT:  Oltersdorf, Tilman
;     TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;     TITLE OF INVENTION:  Acids and Methods of Use
;     NUMBER OF SEQUENCES:  15
;     CORRESPONDENCE ADDRESS:
;       ADDRESSEE:  Campbell and Flores
;       STREET:  4370 La Jolla Village Drive, Suite 700
;       CITY:  San Diego
;       STATE:  California
;       COUNTRY:  United States
;       ZIP:  92122
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy disk
;       COMPUTER:  IBM PC compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/717,123
;       FILING DATE:  20-SEP-1996
;       CLASSIFICATION:  435
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  Campbell, Cathryn A.
;       REGISTRATION NUMBER:  31,815
;       REFERENCE/DOCKET NUMBER:  P-ID 1929
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (619) 535-9001
;       TELEFAX:  (619) 535-8949
;   INFORMATION FOR SEQ ID NO:  3:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  204 amino acids
;       TYPE:  amino acid
;       TOPOLOGY:  linear
US-08-717-123-3

  Query Match             74.6%;  Score 106;  DB 2;  Length 204;
  Best Local Similarity   83.3%;  Pred. No. 5e-09;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||:|||  |||
Db        140 NLWAAQRYGRELRRMTDEFEGSFK 163


RESULT 15
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match             74.6%;  Score 106;  DB 4;  Length 204;
  Best Local Similarity   83.3%;  Pred. No. 5e-09;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||:|||  |||
Db        140 NLWAAQRYGRELRRMTDEFEGSFK 163


Search completed: September 23, 2003, 09:45:40
Job time : 32 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 23, 2003, 09:47:20 ; Search time 389 Seconds
(without alignments)
10.335 Million cell updates/sec

Title: US-09-544-664B-2
Perfect score: 142
Sequence: 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 556269 seqs, 148893369 residues

Total number of hits satisfying chosen parameters: 556269

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	137	96.5	168	9	US-09-922-378-2	Sequence 2, Appli
2	137	96.5	168	9	US-09-894-657-1	Sequence 1, Appli
3	137	96.5	168	9	US-09-894-657-7	Sequence 7, Appli
4	137	96.5	168	12	US-10-209-967-2	Sequence 2, Appli
5	137	96.5	168	14	US-10-066-179-2	Sequence 2, Appli
6	127	89.4	25	15	US-10-059-261-258	Sequence 258, App
7	109	76.8	204	12	US-10-209-967-4	Sequence 4, Appli
8	106	74.6	204	9	US-09-922-378-3	Sequence 3, Appli
9	106	74.6	204	14	US-10-066-179-3	Sequence 3, Appli
10	57	40.1	15	15	US-10-174-105A-147	Sequence 147, App
11	48	33.8	373	12	US-10-168-780-2	Sequence 2, Appli
12	47	33.1	350	11	US-09-940-244-394	Sequence 394, App
13	46.5	32.7	1053	9	US-09-841-132-583	Sequence 583, App
14	46.5	32.7	1053	14	US-10-007-693-97	Sequence 97, Appl
15	46	32.4	35	15	US-10-092-750-1	Sequence 1, Appli
16	46	32.4	113	9	US-09-815-242-13424	Sequence 13424, A
17	46	32.4	138	15	US-10-092-750-241	Sequence 241, App
18	46	32.4	359	15	US-10-136-444-8	Sequence 8, Appli
19	46	32.4	558	12	US-10-342-224-110	Sequence 110, App
20	46	32.4	1270	15	US-10-223-070-8	Sequence 8, Appli
21	45	31.7	404	15	US-10-156-761-10995	Sequence 10995, A
22	45	31.7	441	15	US-10-156-761-13669	Sequence 13669, A
23	45	31.7	454	15	US-10-166-087-46	Sequence 46, Appl
24	45	31.7	541	15	US-10-234-432-37	Sequence 37, Appl
25	45	31.7	662	15	US-10-234-432-75	Sequence 75, Appl
26	45	31.7	678	15	US-10-234-432-38	Sequence 38, Appl
27	45	31.7	901	9	US-09-815-242-11148	Sequence 11148, A
28	44	31.0	272	15	US-10-156-761-11541	Sequence 11541, A
29	44	31.0	362	9	US-09-815-242-10793	Sequence 10793, A
30	44	31.0	414	10	US-09-854-286-18	Sequence 18, Appl
31	44	31.0	426	12	US-10-137-870-218	Sequence 218, App
32	44	31.0	426	12	US-10-140-018-218	Sequence 218, App
33	44	31.0	426	12	US-10-140-021-218	Sequence 218, App
34	44	31.0	426	12	US-10-140-274-218	Sequence 218, App
35	44	31.0	426	12	US-10-140-471-218	Sequence 218, App
36	44	31.0	426	12	US-10-140-807-218	Sequence 218, App
37	44	31.0	426	12	US-10-140-922-218	Sequence 218, App
38	44	31.0	426	12	US-10-140-924-218	Sequence 218, App
39	44	31.0	426	12	US-10-140-926-218	Sequence 218, App
40	44	31.0	426	12	US-10-141-698-218	Sequence 218, App
41	44	31.0	426	12	US-10-141-702-218	Sequence 218, App
42	44	31.0	426	12	US-10-141-704-218	Sequence 218, App
43	44	31.0	426	12	US-10-142-421-218	Sequence 218, App
44	44	31.0	426	12	US-10-142-432-218	Sequence 218, App
45	44	31.0	426	12	US-10-142-767-218	Sequence 218, App

ALIGNMENTS

RESULT 1
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE
; FILE REFERENCE: 480140.428D3
; CURRENT APPLICATION NUMBER: US/09/922,378
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-922-378-2

Query Match 96.5%; Score 137; DB 9; Length 168;
Best Local Similarity 96.3%; Pred. No. 1.2e-12;
Matches 26; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
|||||| ||||||||||||||||||||||
Db 103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129

RESULT 2
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Yue, Henry
; Lal, Preeti
; Shah, Purvi

```
;                 Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: SYNORAB01
;            CLONE: 358673
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             96.5%;  Score 137;  DB 9;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.2e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| |||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 3
US-09-894-657-7
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: GenBank
;            CLONE: 1683637
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match             96.5%;  Score 137;  DB 9;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.2e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| |||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 4
US-10-209-967-2
; Sequence 2, Application US/10209967
; Publication No. US20030171279A1
; GENERAL INFORMATION:
;  APPLICANT: MUNGER, JOSHUA
;  APPLICANT:  ROIZMAN, BERNARD
;  TITLE OF INVENTION: METHODS AND COMPOSITION CONCERNING HERPESVIRUS US3 AND
;  TITLE OF INVENTION:  BAD-INVOLVED APOPTOSIS
;  FILE REFERENCE: ARCD:380US
;  CURRENT APPLICATION NUMBER: US/10/209,967
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/308,929
;  PRIOR FILING DATE: 2001-07-31
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Human BAD
US-10-209-967-2

  Query Match             96.5%;  Score 137;  DB 12;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.2e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| |||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 5
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
```

```
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match             96.5%;  Score 137;  DB 14;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 1.2e-12;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 6
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match             89.4%;  Score 127;  DB 15;  Length 25;
  Best Local Similarity   96.0%;  Pred. No. 4.6e-12;
  Matches   24;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKK 25
              |||||| ||||||||||||||||||||
Db          1 NLWAAQRYGRELRRMSDEFVDSFKK 25


RESULT 7
US-10-209-967-4
; Sequence 4, Application US/10209967
; Publication No. US20030171279A1
; GENERAL INFORMATION:
;  APPLICANT: MUNGER, JOSHUA
;  APPLICANT:  ROIZMAN, BERNARD
;  TITLE OF INVENTION: METHODS AND COMPOSITION CONCERNING HERPESVIRUS US3 AND
;  TITLE OF INVENTION:  BAD-INVOLVED APOPTOSIS
;  FILE REFERENCE: ARCD:380US
;  CURRENT APPLICATION NUMBER: US/10/209,967
;  CURRENT FILING DATE:  2002-07-31
;  PRIOR APPLICATION NUMBER: 60/308,929
;  PRIOR FILING DATE: 2001-07-31
;  NUMBER OF SEQ ID NOS: 25
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-209-967-4

  Query Match             76.8%;  Score 109;  DB 12;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 2.1e-08;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 8
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-922-378-3

  Query Match             74.6%;  Score 106;  DB 9;  Length 204;
  Best Local Similarity   83.3%;  Pred. No. 5.8e-08;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| ||||||||:|||  |||
Db        140 NLWAAQRYGRELRRMTDEFEGSFK 163


RESULT 9
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             74.6%;  Score 106;  DB 14;  Length 204;
  Best Local Similarity   83.3%;  Pred. No. 5.8e-08;
  Matches   20;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| ||||||||:|||  |||
Db        140 NLWAAQRYGRELRRMTDEFEGSFK 163
```

```
RESULT 10
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECIFIC,
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHING
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match             40.1%;  Score 57;  DB 15;  Length 15;
  Best Local Similarity   85.7%;  Pred. No. 0.063;
  Matches   12;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFVDS 22
              ||||||||||||  |
Db          1 GRELRRMSDEFEGS 14


RESULT 11
US-10-168-780-2
; Sequence 2, Application US/10168780
; Publication No. US20030172405A1
; GENERAL INFORMATION:
;  APPLICANT: TANAKA, Hiroshi
;  APPLICANT:  KAYANO, Toshiaki
;  APPLICANT:  YANO, Masahiro
;  APPLICANT:  MATSUOKA, Makoto
;  APPLICANT:  KOBAYASHI, Masatomo
;  TITLE OF INVENTION: GIBBERELLIN 3BETA-HYDROXYLASE GENES OF RICE AND USES THEREOF
;  FILE REFERENCE: SHZ-004US
;  CURRENT APPLICATION NUMBER: US/10/168,780
;  CURRENT FILING DATE:  2002-06-20
;  PRIOR APPLICATION NUMBER: PCT/JP00/09037
;  PRIOR FILING DATE: 2000-12-20
;  PRIOR APPLICATION NUMBER: JP 11/361608
;  PRIOR FILING DATE: 1999-12-20
;  NUMBER OF SEQ ID NOS: 8
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 373
;   TYPE: PRT
;   ORGANISM: Oryza sativa
US-10-168-780-2

  Query Match             33.8%;  Score 48;  DB 12;  Length 373;
  Best Local Similarity   36.4%;  Pred. No. 45;
  Matches    8;  Conservative    8;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          6 QEYGRELRRMSDEFVDSFKKGL 27
              :|: :|:||::|| :  | : |
Db        162 EEFHKEMRRLADELLRLFLRAL 183


RESULT 12
US-09-940-244-394
; Sequence 394, Application US/09940244
; Publication No. US20030044796A1
; GENERAL INFORMATION:
;  APPLICANT: Neri, Bruce P.
;  APPLICANT:  Hall, Jeff G.
;  APPLICANT:  Lyamichev, Victor
;  APPLICANT:  Smith, Lloyd M.
;  TITLE OF INVENTION: Reactions on Dendrimers
;  FILE REFERENCE: FORS-06478
;  CURRENT APPLICATION NUMBER: US/09/940,244
;  CURRENT FILING DATE:  2002-05-06
;  NUMBER OF SEQ ID NOS: 422
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 394
;   LENGTH: 350
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic
US-09-940-244-394

  Query Match             33.1%;  Score 47;  DB 11;  Length 350;
  Best Local Similarity   40.9%;  Pred. No. 58;
  Matches    9;  Conservative    5;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          4 AAQEYGRELRRMSDEFVDSFKK 25
              ||: | :   |::|| |:  ||
Db        119 AARRYAQAAARLTDEMVEDAKK 140


RESULT 13
US-09-841-132-583
; Sequence 583, Application US/09841132
; Patent No. US20020061848A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
;  APPLICANT:  Skeiky, Yasir A.W.
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT AND
;  TITLE OF INVENTION:  DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.469C8
;  CURRENT APPLICATION NUMBER: US/09/841,132
;  CURRENT FILING DATE:  2001-04-23
;  NUMBER OF SEQ ID NOS: 599
;  SOFTWARE: FastSEQ for Windows Version 3.0/4.0
; SEQ ID NO 583
;   LENGTH: 1053
;   TYPE: PRT
;   ORGANISM: C. Trachomatis D serovar
US-09-841-132-583

  Query Match             32.7%;  Score 46.5;  DB 9;  Length 1053;
  Best Local Similarity   40.7%;  Pred. No. 2.3e+02;
  Matches   11;  Conservative    7;  Mismatches    6;  Indels    3;  Gaps    1;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              ||: ||  |::|  |   :: ::||||
Db        969 NLYLAQPDGKKLSNM---YLTAWKKGL 992


RESULT 14
US-10-007-693-97
; Sequence 97, Application US/10007693
; Publication No. US20020146776A1
; GENERAL INFORMATION:
;  APPLICANT: Bhatia, Ajay
```

```
;  APPLICANT:  Probst, Peter
;  TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
;  TITLE OF INVENTION:  AND DIAGNOSIS OF CHLAMYDIAL INFECTION
;  FILE REFERENCE: 210121.515C2
;  CURRENT APPLICATION NUMBER: US/10/007,693
;  CURRENT FILING DATE:  2001-12-05
;  NUMBER OF SEQ ID NOS: 157
; SEQ ID NO 97
;   LENGTH: 1053
;   TYPE: PRT
;   ORGANISM: Chlamydia trachomatis serovar D
US-10-007-693-97

  Query Match             32.7%;  Score 46.5;  DB 14;  Length 1053;
  Best Local Similarity   40.7%;  Pred. No. 2.3e+02;
  Matches   11;  Conservative    7;  Mismatches    6;  Indels    3;  Gaps    1;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              ||: ||  |:::|  |   :: ::||||
Db        969 NLYLAQPDGKKLSNM---YLTAWKKGL 992


RESULT 15
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-1

  Query Match             32.4%;  Score 46;  DB 15;  Length 35;
  Best Local Similarity   55.6%;  Pred. No. 6.8;
  Matches   10;  Conservative    2;  Mismatches    2;  Indels    4;  Gaps    1;

Qy          2 LWAAQEYGRELRRMSDEF 19
              :| ||    ||||: |||
Db         15 IWIAQ----ELRRIGDEF 28


Search completed: September 23, 2003, 10:03:35
Job time : 390 secs
```

OM protein - protein search, using sw model

Run on: September 23, 2003, 09:43:16 ; Search time 22 Seconds
(without alignments)
57.715 Million cell updates/sec

Title: US-09-544-664B-2
Perfect score: 142
Sequence: 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	137	96.5	168	1	BAD_HUMAN	Q92934 homo sapien
2	109	76.8	204	1	BAD_MOUSE	Q61337 mus musculu
3	109	76.8	205	1	BAD_RAT	O35147 rattus norv
4	51	35.9	653	1	HT2A_HUMAN	Q13049 homo sapien
5	50.5	35.6	503	1	MTBR_BPRH1	P09915 bacteriopha
6	50.5	35.6	521	1	PPCK_THETN	Q8r943 thermoanaer
7	50	35.2	370	1	AROG_YEAST	P32449 saccharomyc
8	49.5	34.9	320	1	BXD1_DROME	Q9veb3 drosophila
9	49	34.5	478	1	HXK_DEBOC	P50506 debaryomyce
10	48	33.8	370	1	AROG_CANAL	P79023 candida alb
11	48	33.8	526	1	VP5_BTV2A	P30209 bluetongue
12	48	33.8	629	1	SYM_THEMA	O33925 thermotoga
13	47.5	33.5	537	1	PPCK_RHISN	P43086 rhizobium s
14	47.5	33.5	595	1	VG13_BPMD2	O64206 mycobacteri
15	47	33.1	631	1	RPSD_BORBU	P52323 borrelia bu
16	47	33.1	851	1	CE05_MOUSE	Q8k2h3 mus musculu
17	46.5	32.7	181	1	IPYR_MYCPE	Q8evw9 mycoplasma
18	46.5	32.7	333	1	Y591_AQUAE	O66853 aquifex aeo
19	46.5	32.7	1047	1	RIR1_CHLMU	Q9p193 chlamydia m
20	46.5	32.7	1047	1	RIR1_CHLTR	O84834 chlamydia t
21	46	32.4	198	1	BIM_HUMAN	O43521 homo sapien
22	46	32.4	306	1	MRAW_RICPR	Q9zcy2 rickettsia
23	46	32.4	348	1	FLS_PETHY	Q07512 petunia hyb
24	46	32.4	467	1	TNAA_AERPE	Q9yci2 aeropyrum p
25	46	32.4	526	1	VP5_BTV1A	P12436 bluetongue
26	46	32.4	1513	1	STU1_YEAST	P38198 saccharomyc
27	45.5	32.0	143	1	YXAD_BACSU	P42103 bacillus su
28	45.5	32.0	527	1	PPCK_BACHD	Q9k7q7 bacillus ha
29	45.5	32.0	659	1	YYBT_BACSU	P37484 bacillus su
30	45	31.7	878	1	SYV_METJA	Q58413 methanococc
31	45	31.7	901	1	SECA_HAEIN	P43803 haemophilus
32	44.5	31.3	382	1	ISC1_ARCFU	O30052 archaeoglob
33	44.5	31.3	1339	1	DPOA_LEIDO	O00874 leishmania
34	44	31.0	125	1	VSIS_REOVJ	P32868 reovirus (t
35	44	31.0	232	1	AP3_ARATH	P35632 arabidopsis
36	44	31.0	261	1	YGFG_ECOLI	P52045 escherichia
37	44	31.0	310	1	ACP2_ENTHI	P36185 entamoeba h
38	44	31.0	311	1	ALSB_ECOLI	P39265 escherichia
39	44	31.0	315	1	CPP2_ENTHI	Q01958 entamoeba h
40	44	31.0	351	1	E2BB_RABIT	Q28690 oryctolagus
41	44	31.0	422	1	PAN1_HUMAN	Q96rd7 homo sapien
42	44	31.0	453	1	RMUC_PSEAE	Q9i4u3 pseudomonas
43	44	31.0	534	1	CN9A_MOUSE	O70628 mus musculu
44	44	31.0	661	1	HT31_ARATH	Q04996 arabidopsis
45	44	31.0	704	1	KPC1_CAEEL	P34722 caenorhabdi

ALIGNMENTS

RESULT 1
BAD_HUMAN
ID BAD_HUMAN STANDARD; PRT; 168 AA.
AC Q92934; O14803;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN BAD OR BBC6 OR BCL2L8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT "A human protein that interacts with Bcl-2 and have homology to mouse
RT BAD.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX MEDLINE=97083574; PubMed=8929532;
RA Wang H.-G., Rapp U.R., Reed J.C.;
RT "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL Cell 87:629-638(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Takayama S., Reed J.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND DIMERIZATION.
RC TISSUE=Bone marrow;
RX MEDLINE=98049554; PubMed=9388232;
RA Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT "Dimerization properties of human BAD.";
RL J. Biol. Chem. 272:30866-30872(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match             96.5%;  Score 137;  DB 1;  Length 168;
  Best Local Similarity   96.3%;  Pred. No. 8.2e-13;
  Matches   26;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              |||||| ||||||||||||||||||||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGL 129


RESULT 2
BAD_MOUSE
ID   BAD_MOUSE      STANDARD;      PRT;   204 AA.
AC   Q61337;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD OR BBC6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Thymus;
RX   MEDLINE=95136361; PubMed=7834748;
RA   Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT   "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT   promotes cell death.";
RL   Cell 80:285-291(1995).
RN   [2]
RP   PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX   MEDLINE=98022383; PubMed=9381178;
RA   Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT   "Interleukin-3-induced phosphorylation of BAD through the protein
RT   kinase Akt.";
RL   Science 278:687-689(1997).
RN   [3]
RP   MUTAGENESIS OF SERINE RESIDUES.
RX   MEDLINE=20403302; PubMed=10949026;
RA   Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA   Greenberg M.E.;
RT   "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT   BH3 domain phosphorylation.";
RL   Mol. Cell 6:41-51(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC       Appears to act as a link between growth factor receptor signaling
CC       and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-155, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
```

```
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             76.8%;  Score 109;  DB 1;  Length 204;
  Best Local Similarity   87.5%;  Pred. No. 1.2e-08;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| ||||||||||||  |||
Db        140 NLWAAQRYGRELRRMSDEFEGSFK 163


RESULT 3
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;

  Query Match             76.8%;  Score 109;  DB 1;  Length 205;
```

```
  Best Local Similarity   87.5%;  Pred. No. 1.2e-08;
  Matches   21;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFK 24
              |||||| |||||||||||  |||
Db        141 NLWAAQRYGRELRRMSDEFEGSFK 164


RESULT 4
HT2A_HUMAN
ID   HT2A_HUMAN     STANDARD;      PRT;   653 AA.
AC   Q13049; Q9NQP8;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc-finger protein HT2A (72 kDa Tat-interacting protein) (Tripartite
DE   motif-containing protein 32).
GN   TRIM32 OR HT2A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95297135; PubMed=7778269;
RA   Fridell R.A., Harding L.S., Bogerd H.P., Cullen B.R.;
RT   "Identification of a novel human zinc finger protein that
RT   specifically interacts with the activation domain of lentiviral Tat
RT   proteins.";
RL   Virology 209:347-357(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sehra H.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: MAY PLAY A SIGNIFICANT ROLE IN MEDIATING THE BIOLOGICAL
CC       ACTIVITY OF THE HIV-1 TAT PROTEIN IN VIVO. BINDS SPECIFICALLY TO
CC       THE ACTIVATION DOMAIN OF HIV-1 TAT AND CAN ALSO INTERACT WITH THE
CC       HIV-2 AND EIAV TAT PROTEINS IN VIVO.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: SPLEEN, THYMUS, PROSTRATE, TESTIS, OVARY,
CC       INTESTINE AND COLON.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC   -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18543; AAA86474.1; -.
DR   EMBL; AL133284; CAB92723.1; -.
DR   EMBL; BC003154; AAH03154.1; -.
DR   HSSP; P29590; 1BOR.
DR   Genew; HGNC:16380; TRIM32.
DR   MIM; 602290; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   InterPro; IPR001258; NHL.
DR   InterPro; IPR000315; Znf_Bbox.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF01436; NHL; 5.
DR   Pfam; PF00643; zf-B_box; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00336; BBOX; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Zinc-finger; Nuclear protein.
FT   DOMAIN        2      6       POLY-ALA.
FT   ZN_FING      20     65       RING-TYPE.
FT   ZN_FING     103    133       B BOX-TYPE.
FT   CONFLICT     27     27       F -> I (IN REF. 1).
SQ   SEQUENCE   653 AA;  71988 MW;  D83B1595CA8378FD CRC64;

  Query Match             35.9%;  Score 51;  DB 1;  Length 653;
  Best Local Similarity   57.9%;  Pred. No. 11;
  Matches   11;  Conservative    1;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 QEYGRELRRMSDEFVDSFK 24
              |||| | ||: ||    | |
Db        186 QEYGHEERRVQDELARSRK 204


RESULT 5
MTBR_BPRH1
ID   MTBR_BPRH1     STANDARD;      PRT;   503 AA.
AC   P09915;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Modification methylase RholIsI (EC 2.1.1.73) (Cytosine-specific
DE   methyltransferase RholIsI) (Bsu PIIs) (M.RholIsI).
OS   Bacteriophage rho-11s.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC   Lambda-like viruses.
OX   NCBI_TaxID=10735;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87246516; PubMed=3109889;
RA   Behrens B., Noyer-Weidner M., Pawlek B., Lauster R., Balganesh T.S.,
RA   Trautner T.A.;
RT   "Organization of multispecific DNA methyltransferases encoded by
RT   temperate Bacillus subtilis phages.";
RL   EMBO J. 6:1137-1142(1987).
RN   [2]
RP   REVISION TO 476.
RA   Trautner T.A.;
RL   Submitted (SEP-1987) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS ENZYME METHYLATES CYTOSINE WITHIN THE SEQUENCES
CC       GGCC AND GAGCTC.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA cytosine = S-
CC       adenosyl-L-homocysteine + DNA 5-methylcytosine.
CC   -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
```

```
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X05242; CAA28869.1; -.
DR   PIR; A28137; CTBPRH.
DR   HSSP; P05102; 6MHT.
DR   REBASE; 2835; M.RhollsI.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   Pfam; PF00145; DNA_methylase; 1.
DR   PRINTS; PR00105; C5METTRFRASE.
DR   TIGRFAMs; TIGR00675; dcm; 1.
DR   PROSITE; PS00094; C5_MTASE_1; 1.
DR   PROSITE; PS00095; C5_MTASE_2; 1.
KW   Transferase; Methyltransferase; Restriction system.
FT   ACT_SITE     78      78       BY SIMILARITY.
SQ   SEQUENCE   503 AA;  57129 MW;  AAAFB8FE01B8129E CRC64;

  Query Match             35.6%;  Score 50.5;  DB 1;  Length 503;
  Best Local Similarity   50.0%;  Pred. No. 9.5;
  Matches   10;  Conservative    5;  Mismatches    4;  Indels    1;  Gaps    1;

Qy          3 WAAQE-YGRELRRMSDEFVD 21
              |:||:  || || : :|:||
Db        207 WSAQDIVGRRLREILEEYVD 226


RESULT 6
PPCK_THETN
ID   PPCK_THETN     STANDARD;      PRT;   521 AA.
AC   Q8R943;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP
DE   carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN   PCKA OR TTE1783.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC   Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4 / JCM 11007;
RX   MEDLINE=21992816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
CC   -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
CC       + CO(2).
CC   -!- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]
CC       family.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AE013131; AAM24977.1; -.
DR   HAMAP; MF_00453; -; 1.
DR   InterPro; IPR001272; PEPCK_ATP.
DR   Pfam; PF01293; PEPCK_ATP; 1.
DR   ProDom; PD004723; PEPCK_ATP; 1.
DR   TIGRFAMs; TIGR00224; pckA; 1.
DR   PROSITE; PS00532; PEPCK_ATP; 1.
KW   Gluconeogenesis; Lyase; Decarboxylase; ATP-binding; Complete proteome.
FT   NP_BIND     227    234       ATP (BY SIMILARITY).
SQ   SEQUENCE   521 AA;  58771 MW;  1783A3320B106341 CRC64;

  Query Match             35.6%;  Score 50.5;  DB 1;  Length 521;
  Best Local Similarity   30.8%;  Pred. No. 9.9;
  Matches    8;  Conservative    9;  Mismatches    8;  Indels    1;  Gaps    1;

Qy          1 NLWA-AQEYGRELRRMSDEFVDSFKK 25
                | |   :|| :  :|::  |:::|:|
Db        481 NTWKDKEEYDKTAKRLAQRFIENFQK 506


RESULT 7
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between  the Swiss Institute of Bioinformatics  and the  EMBL outstation -
CC   the European Bioinformatics Institute.  There are no  restrictions on  its
CC   use  by  non-profit  institutions as long  as its content  is  in  no  way
CC   modified and this statement is not removed.  Usage  by  and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match           35.2%;  Score 50;  DB 1;  Length 370;
  Best Local Similarity   55.6%;  Pred. No. 8.1;
  Matches   10;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 NLWAAQEYGRELRRMSDE 18
              :| ||||| |:::|||
Db         80 DLEAAQEYALRLKKLSDE 97


RESULT 8
BXD1_DROME
ID   BXD1_DROME     STANDARD;      PRT;   320 AA.
AC   Q9VEB3; Q95U84;
DT   16-OCT-2001 (Rel. 40, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Brix domain containing protein 1 homolog.
GN   CG7993.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=22426066; PubMed=12537569;
RA   Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA   George R.A., Guarin H., Kronmiller B., Pacleb J.M., Park S., Wan K.H.,
RA   Rubin G.M., Celniker S.E.;
RT   "A Drosophila full-length cDNA resource.";
RL   Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC   -!- SUBCELLULAR LOCATION: Nuclear; nucleolar (By similarity).
CC   -!- SIMILARITY: Contains 1 Brix domain.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to erroneous
CC       gene model prediction.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE003721; AAF55514.1; ALT_SEQ.
DR   EMBL; AY058248; AAL13477.1; -.
DR   FlyBase; FBgn0038585; CG7993.
DR   InterPro; IPR007109; Brix.
DR   Pfam; PF04427; Brix; 1.
KW   Hypothetical protein; Nuclear protein.
FT   DOMAIN       30    243       BRIX.
FT   CONFLICT    188    188       I -> T (IN REF. 1).
SQ   SEQUENCE   320 AA;  36509 MW;  EE98936DD68B3703 CRC64;

  Query Match           34.9%;  Score 49.5;  DB 1;  Length 320;
  Best Local Similarity   43.5%;  Pred. No. 8.1;
  Matches   10;  Conservative    6;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          3 WAAQEYGRELRRMSDEFVDSFKK 25
              ||  |   ||||: : |:|:|::
Db        149 WAQTE---ELRRLRNLFIDTFQR 168


RESULT 9
HXK_DEBOC
ID   HXK_DEBOC      STANDARD;      PRT;   478 AA.
AC   P50506;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-JUL-1998 (Rel. 36, Last annotation update)
DE   Hexokinase (EC 2.7.1.1).
GN   HXK.
OS   Debaryomyces occidentalis (Yeast) (Schwanniomyces occidentalis).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Debaryomyces.
OX   NCBI_TaxID=27300;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 2322 / CBS 819;
RX   MEDLINE=95339407; PubMed=7614556;
RA   Rose M.;
RT   "Molecular and biochemical characterization of the hexokinase from
RT   the starch-utilizing yeast Schwanniomyces occidentalis.";
RL   Curr. Genet. 27:330-338(1995).
CC   -!- CATALYTIC ACTIVITY: ATP + D-hexose = ADP + D-hexose 6-phosphate.
CC   -!- PATHWAY: FIRST STEP OF SEVERAL METABOLIC PATHWAYS.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE HEXOKINASE FAMILY.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; S78714; AAB34892.1; -.
DR   PIR; S57203; S57203.
DR   HSSP; Q26609; 1BDG.
DR   InterPro; IPR001312; Hexokinase.
DR   Pfam; PF03727; hexokinase2; 1.
DR   Pfam; PF00349; hexokinase; 1.
DR   PRINTS; PR00475; HEXOKINASE.
DR   ProDom; PD001109; Hexokinase; 1.
DR   PROSITE; PS00378; HEXOKINASES; 1.
KW   Transferase; Kinase; Glycolysis; Allosteric enzyme; ATP-binding.
FT   BINDING     111    111       ATP (BY SIMILARITY).
FT   DOMAIN      151    177       GLUCOSE-BINDING (POTENTIAL).
SQ   SEQUENCE   478 AA;  53066 MW;  080D5F9134478ABA CRC64;

  Query Match             34.5%;  Score 49;  DB 1;  Length 478;
  Best Local Similarity   42.1%;  Pred. No. 15;
  Matches    8;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFVDSFKKGL 27
              |  ||:::| |:   :|||
Db         38 GETLRKITDHFISELEKGL 56


RESULT 10
AROG_CANAL
ID   AROG_CANAL     STANDARD;      PRT;   370 AA.
AC   P79023;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 11651 / B792;
RA   Sousa S., Pereira S.A., Livi G.P.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   PARTIAL SEQUENCE FROM N.A.
RX   MEDLINE=96207468; PubMed=8625423;
RA   Pereira S.A., Livi G.P.;
RT   "Aromatic amino-acid biosynthesis in Candida albicans: identification
RT   of the ARO4 gene encoding a second DAHP synthase.";
RL   Curr. Genet. 29:441-445(1996).
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE (BY SIMILARITY).
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U53216; AAB48240.1; -.
DR   HSSP; P00886; 1QR7.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  40291 MW;  11E5E324C8D7B6DB CRC64;

  Query Match             33.8%;  Score 48;  DB 1;  Length 370;
  Best Local Similarity   47.1%;  Pred. No. 16;
  Matches    8;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          5 AQEYGRELRRMSDEFVD 21
              | |||: |::::||  |
Db         74 ALEYGKRLKKLADELKD 90


RESULT 11
VP5_BTV2A
ID   VP5_BTV2A      STANDARD;      PRT;   526 AA.
AC   P30209;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Outer capsid protein VP5.
GN   S6.
OS   Bluetongue virus (serotype 2 / isolate USA).
OC   Viruses; dsRNA viruses; Reoviridae; Orbivirus.
OX   NCBI_TaxID=10907;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90209358; PubMed=2157314;
RA   Hirasawa T., Roy P.;
RT   "The complete nucleotide sequence of VP5 of a strain of bluetongue
RT   virus of serotype 2 isolated in the USA reveals its close
RT   relationship with a virus of serotype 1 isolated in Australia.";
RL   Virus Res. 15:107-112(1990).
CC   -!- FUNCTION: THE VP5 PROTEIN IS ONE OF THE TWO PROTEINS (WITH VP2)
CC       WHICH CONSTITUTE THE VIRUS PARTICLE OUTER CAPSID.
CC   -!- SIMILARITY: BELONGS TO THE REOVIRUSES VP5 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X62283; CAA44172.1; -.
DR   PIR; A43486; P5XRBU.
DR   InterPro; IPR000145; Orbi_VP5.
DR   Pfam; PF00901; Orbi_VP5; 1.
KW   Coat protein.
SQ   SEQUENCE   526 AA;  58953 MW;  DE50D6013B983A04 CRC64;

  Query Match             33.8%;  Score 48;  DB 1;  Length 526;
  Best Local Similarity   41.7%;  Pred. No. 23;
  Matches   10;  Conservative    4;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          4 AAQEYGRELRRMSDEFVDSFKKGL 27
              | ||  :||  : || :  ||| :
Db        309 AIQENHKELMHIKDEILPRFKKAM 332


RESULT 12
```

```
SYM_THEMA
ID   SYM_THEMA      STANDARD;      PRT;   629 AA.
AC   O33925;
DT   15-JUL-1999 (Rel. 38, Created)
DT   15-JUL-1999 (Rel. 38, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Methionyl-tRNA synthetase (EC 6.1.1.10) (Methionine--tRNA ligase)
DE   (MetRS).
GN   METG OR METS OR TM1085.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=98021416; PubMed=9380682;
RA   Guipaud O., Marguet E., Noll K.M., de la Tour C.B., Forterre P.;
RT   "Both DNA gyrase and reverse gyrase are present in the
RT   hyperthermophilic bacterium Thermotoga maritima.";
RL   Proc. Natl. Acad. Sci. U.S.A. 94:10606-10611(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: Is required not only for elongation of protein synthesis
CC       but also for the initiation of all mRNA translation through
CC       initiator tRNA(fMet) aminoacylation (By similarity).
CC   -!- CATALYTIC ACTIVITY: ATP + L-methionine + tRNA(Met) = AMP +
CC       diphosphate + L-methionyl-tRNA(Met).
CC   -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to class-I aminoacyl-tRNA synthetase family.
CC       MetG subfamily 2A.
CC   -!- SIMILARITY: Contains 1 tRNA-binding domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U76417; AAB87143.1; -.
DR   EMBL; AE001768; AAD36162.1; -.
DR   PIR; E72297; E72297.
DR   HSSP; P23395; 1A8H.
DR   TIGR; TM1085; -.
DR   HAMAP; MF_01228; fused; 1.
DR   InterPro; IPR004495; MetG_Cterm.
DR   InterPro; IPR002300; tRNA-synt_1a.
DR   InterPro; IPR001412; tRNA-synt_I.
DR   InterPro; IPR002304; tRNA-synt_met.
DR   InterPro; IPR002547; tRNA_bind.
DR   Pfam; PF00133; tRNA-synt_1; 1.
DR   Pfam; PF01588; tRNA_bind; 1.
DR   PIRSF; PIRSF001528; MetRS_dimerising; 1.
DR   PRINTS; PR01041; TRNASYNTHMET.
DR   TIGRFAMs; TIGR00398; metG; 1.
DR   TIGRFAMs; TIGR00399; metG_C_term; 1.
DR   PROSITE; PS00178; AA_TRNA_LIGASE_I; FALSE_NEG.
DR   PROSITE; PS50886; TRBD; 1.
KW   Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW   RNA-binding; tRNA-binding; Metal-binding; Zinc; Complete proteome.
FT   SITE        10     20       "HIGH" REGION.
FT   SITE       297    301       "KMSKS" REGION.
FT   DOMAIN     529    629       TRNA-BINDING.
FT   METAL      125    125       ZINC (BY SIMILARITY).
FT   METAL      128    128       ZINC (BY SIMILARITY).
FT   METAL      146    146       ZINC (BY SIMILARITY).
FT   METAL      149    149       ZINC (BY SIMILARITY).
FT   BINDING    300    300       ATP (BY SIMILARITY).
SQ   SEQUENCE   629 AA;  73004 MW;  B0E0759F7C78ACEC CRC64;

  Query Match             33.8%;  Score 48;  DB 1;  Length 629;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches    9;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          2 LWAAQEYGRELRRMSDEFVDSFKK 25
              | |||: |:: :   ||  : ||:
Db         55 LQAAQQAGKDPQEFCDELAEKFKR 78


RESULT 13
PPCK_RHISN
ID   PPCK_RHISN     STANDARD;      PRT;   537 AA.
AC   P43086;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phosphoenolpyruvate carboxykinase [ATP] (EC 4.1.1.49) (PEP
DE   carboxykinase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN   PCKA.
OS   Rhizobium sp. (strain NGR234).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX   NCBI_TaxID=394;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92079905; PubMed=1720862;
RA   Oesteraas M., Finan T.M., Stanley J.;
RT   "Site-directed mutagenesis and DNA sequence of pckA of Rhizobium
RT   NGR234, encoding phosphoenolpyruvate carboxykinase: gluconeogenesis
RT   and host-dependent symbiotic phenotype.";
RL   Mol. Gen. Genet. 230:257-269(1991).
CC   -!- CATALYTIC ACTIVITY: ATP + oxaloacetate = ADP + phosphoenolpyruvate
CC       + CO(2).
CC   -!- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: Belongs to the phosphoenolpyruvate carboxykinase [ATP]
CC       family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X63291; CAA44925.1; -.
DR   HSSP; P22259; 1AQ2.
DR   HAMAP; MF_00453; -; 1.
DR   InterPro; IPR001272; PEPCK_ATP.
DR   Pfam; PF01293; PEPCK_ATP; 1.
DR   ProDom; PD004723; PEPCK_ATP; 1.
DR   TIGRFAMs; TIGR00224; pckA; 1.
DR   PROSITE; PS00532; PEPCK_ATP; 1.
KW   Gluconeogenesis; Lyase; Decarboxylase; ATP-binding.
FT   NP_BIND     236    243       ATP (BY SIMILARITY).
SQ   SEQUENCE   537 AA;  58370 MW;  B2CED7FA54326B1E CRC64;

  Query Match             33.5%;  Score 47.5;  DB 1;  Length 537;
  Best Local Similarity   41.7%;  Pred. No. 28;
  Matches   10;  Conservative    5;  Mismatches    8;  Indels    1;  Gaps    1;
```

```
Qy          3 WAA-QEYGRELRRMSDEFVDSFKK 25
              |:| | |  : ||: | |: :| |
Db        493 WSAGQAYDAQARRLVDMFIANFAK 516


RESULT 14
VG13_BPMD2
ID   VG13_BPMD2     STANDARD;      PRT;   595 AA.
AC   O64206;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gene 13 protein (GP13).
GN   13.
OS   Mycobacteriophage D29.
OC   Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX   NCBI_TaxID=28369;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98300335; PubMed=9636706;
RA   Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT   "Genome structure of mycobacteriophage D29: implications for phage
RT   evolution.";
RL   J. Mol. Biol. 279:143-164(1998).
CC   -!- SIMILARITY: BELONGS TO THE PHAGE TERMINASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF022214; AAC18453.1; -.
DR   PIR; B72801; B72801.
DR   InterPro; IPR005021; Phage_termin.
DR   Pfam; PF03354; Phage_terminase; 1.
SQ   SEQUENCE   595 AA;  66397 MW;  AFD123ED5371E263 CRC64;

  Query Match             33.5%;  Score 47.5;  DB 1;  Length 595;
  Best Local Similarity   45.5%;  Pred. No. 31;
  Matches   10;  Conservative    5;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          1 NLWAAQEYGRELRRMSDEFVDS 22
              ::|  |:|| |: |   ||||:
Db        435 DIWDPQKYGGEVPR---EFVDA 453


RESULT 15
RPSD_BORBU
ID   RPSD_BORBU     STANDARD;      PRT;   631 AA.
AC   P52323;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RNA polymerase sigma factor rpoD (Sigma-70).
GN   RPOD OR BB0712.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE OF 89-631 FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RA   Pan M.;
RL   Thesis (1994), National Taiwan University, Taiwan.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
RN   [3]
RP   SEQUENCE OF 165-614 FROM N.A.
RC   STRAIN=297;
RA   Pan M., Yeh J., Tsai C.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC       ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC       THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC       BACTERIA.
CC   -!- SIMILARITY: Belongs to the sigma-70 factor family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U17591; AAC44104.1; -.
DR   EMBL; AE001171; AAC67061.1; -.
DR   EMBL; U68006; AAC45100.1; -.
DR   PIR; G70188; G70188.
DR   HSSP; P00579; 1SIG.
DR   TIGR; BB0712; -.
DR   InterPro; IPR000943; Sigma_70.
DR   Pfam; PF03979; sigma70_r1_1; 1.
DR   Pfam; PF00140; sigma70_r1_2; 1.
DR   Pfam; PF04542; sigma70_r2; 1.
DR   Pfam; PF04539; sigma70_r3; 1.
DR   Pfam; PF04545; sigma70_r4; 1.
DR   PROSITE; PS00715; SIGMA70_1; 1.
DR   PROSITE; PS00716; SIGMA70_2; 1.
KW   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Complete proteome.
FT   DOMAIN      419    432       POLYMERASE CORE BINDING (POTENTIAL).
FT   DNA_BIND    589    608       H-T-H MOTIF (BY SIMILARITY).
SQ   SEQUENCE   631 AA;  73642 MW;  BD565AB7D8F44796 CRC64;

  Query Match             33.1%;  Score 47;  DB 1;  Length 631;
  Best Local Similarity   50.0%;  Pred. No. 39;
  Matches    9;  Conservative    3;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          8 YGRELRRMSDEFVDSFKK 25
              |  |||  ||:::||  |
Db        279 YQEELRIFSDDYIDSANK 296


Search completed: September 23, 2003, 09:47:46
Job time : 26 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 23, 2003, 09:43:16 ; Search time 96 Seconds
(without alignments)
72.577 Million cell updates/sec

Title: US-09-544-664B-2
Perfect score: 142
Sequence: 1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	59.2	146	13	Q9I9N2	Q9i9n2 brachydanio
2	54	38.0	606	10	Q8H4M3	Q8h4m3 oryza sativ
3	53	37.3	85	16	Q8R8Y1	Q8r8y1 thermoanaer
4	52	36.6	889	10	Q8LMF2	Q8lmf2 oryza sativ
5	51	35.9	335	10	O64692	O64692 arabidopsis
6	51	35.9	722	17	Q8TVX8	Q8tvx8 methanopyru
7	51	35.9	1248	16	Q9HZQ3	Q9hzq3 pseudomonas
8	50	35.2	173	4	Q8N955	Q8n955 homo sapien
9	50	35.2	230	16	Q8XXS6	Q8xxs6 ralstonia s
10	50	35.2	330	16	Q8F5R1	Q8f5r1 leptospira
11	50	35.2	363	16	Q8EH69	Q8eh69 shewanella
12	50	35.2	371	5	Q9U1L1	Q9u1l1 drosophila
13	50	35.2	389	5	Q9W5B6	Q9w5b6 drosophila
14	49.5	34.9	320	5	Q95U84	Q95u84 drosophila
15	49	34.5	214	16	Q8KDC3	Q8kdc3 chlorobium
16	49	34.5	247	2	Q9AJL5	Q9ajl5 streptomyce
17	49	34.5	378	17	Q8PYK0	Q8pyk0 methanosarc
18	49	34.5	564	16	Q9RUK9	Q9ruk9 deinococcus
19	49	34.5	1039	17	Q8TME7	Q8tme7 methanosarc
20	49	34.5	1350	5	O17342	O17342 caenorhabdi
21	49	34.5	1690	17	Q9HKL0	Q9hkl0 thermoplasm
22	49	34.5	1713	17	Q97B15	Q97b15 thermoplasm
23	48.5	34.2	130	16	Q8R937	Q8r937 thermoanaer
24	48.5	34.2	937	16	Q97H37	Q97h37 clostridium
25	48.5	34.2	1303	12	Q9QBU0	Q9qbu0 potato mop-
26	48	33.8	196	16	Q8VJS3	Q8vjs3 mycobacteri
27	48	33.8	223	16	Q10843	Q10843 mycobacteri
28	48	33.8	288	10	Q94ID5	Q94id5 oryza sativ
29	48	33.8	295	16	Q9HY62	Q9hy62 pseudomonas
30	48	33.8	339	10	Q9FQ05	Q9fq05 atrichum an
31	48	33.8	357	16	Q8PPY8	Q8ppy8 xanthomonas
32	48	33.8	370	10	Q94ID4	Q94id4 oryza sativ
33	48	33.8	373	10	Q9FU53	Q9fu53 oryza sativ
34	48	33.8	373	10	Q94IE4	Q94ie4 oryza sativ
35	48	33.8	526	12	Q8BAD7	Q8bad7 bluetongue
36	48	33.8	526	12	Q8BAD6	Q8bad6 bluetongue
37	48	33.8	709	10	Q8GUI8	Q8gui8 arabidopsis
38	47.5	33.5	233	17	O29007	O29007 archaeoglob
39	47.5	33.5	234	16	Q8ZAP2	Q8zap2 yersinia pe
40	47.5	33.5	238	16	Q99UU3	Q99uu3 staphylococ
41	47.5	33.5	279	10	Q8W4Z1	Q8w4z1 brassica ol
42	47.5	33.5	309	10	Q9SXJ6	Q9sxj6 arabidopsis
43	47.5	33.5	309	10	Q94JY6	Q94jy6 arabidopsis
44	47.5	33.5	310	10	O82421	O82421 arabidopsis
45	47.5	33.5	311	10	Q9ST19	Q9st19 brassica ca

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 59.2%; Score 84; DB 13; Length 146;
Best Local Similarity 62.5%; Pred. No. 0.00023;
Matches 15; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 LWAAQEYGRELRRMSDEFVDSFKK 25
              |||::||::||||||||    |:
Db         89 LWAAKKYGQQLRRMSDEFDKGMKR 112
```

RESULT 2
Q8H4M3
ID Q8H4M3 PRELIMINARY; PRT; 606 AA.
AC Q8H4M3;

```
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   OJ1116_C08.8 protein.
GN   OJ1116_C08.8.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Sasaki T., Matsumoto T., Yamamoto K.;
RT   "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 7, BAC
RT   clone:OJ1116_C08.";
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AP004002; BAC20651.1; -.
SQ   SEQUENCE   606 AA;  68758 MW;  D453B0DEF5AE2D94 CRC64;

  Query Match             38.0%;  Score 54;  DB 10;  Length 606;
  Best Local Similarity   40.0%;  Pred. No. 26;
  Matches   10;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          3 WAAQEYGRELRRMSDEFVDSFKKGL 27
              || :|   |:||: : :::: ||||
Db        305 WALKEAPWEMRRLHNFYMEASKKGL 329


RESULT 3
Q8R8Y1
ID   Q8R8Y1      PRELIMINARY;      PRT;    85 AA.
AC   Q8R8Y1;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein TTE1854.
GN   TTE1854.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC   Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4 / JCM 11007;
RX   MEDLINE=21992816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
DR   EMBL; AE013137; AAM25042.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   85 AA;  10314 MW;  07531376980DB1F1 CRC64;

  Query Match             37.3%;  Score 53;  DB 16;  Length 85;
  Best Local Similarity   42.9%;  Pred. No. 4;
  Matches    9;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 QEYGRELRRMSDEFVDSFKKG 26
              ||:   |: :||  :: ||||
Db          9 QEFNERLKEVSDNILEFFKKG 29


RESULT 4
Q8LMF2
ID   Q8LMF2      PRELIMINARY;      PRT;   889 AA.
AC   Q8LMF2;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   OJ1325D05.18.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   McCombie W.R., de la Bastide M., Spiegel L., Preston R., Ferraro K.,
RA   Kuit K., Nascimento L., Zutavern T., Balija V., Bell M., Baker J.,
RA   Santos L., Miller B., Katzenberger F., Muller S., King L., Yang C.,
RA   Dike S., O'Shaughnessy A., Palmer L., Dedhia N.;
RT   "Genomic sequence for Oryza sativa, Nipponbare strain, clone
RT   OJ1325D05, from chromosome 10, complete sequence.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC115686; AAM74351.1; -.
DR   Gramene; Q8LMF2; -.
KW   Hypothetical protein.
SQ   SEQUENCE   889 AA;  101583 MW;  C47D8715883D6376 CRC64;

  Query Match             36.6%;  Score 52;  DB 10;  Length 889;
  Best Local Similarity   40.0%;  Pred. No. 79;
  Matches   10;  Conservative    6;  Mismatches    9;  Indels    0;  Gaps    0;

Qy          3 WAAQEYGRELRRMSDEFVDSFKKGL 27
              ||  |   |::|:   ::|: ||||
Db        627 WALDEAPWEMKRLHTWYMDASKKGL 651


RESULT 5
O64692
ID   O64692      PRELIMINARY;      PRT;   335 AA.
AC   O64692;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative GA4 protein.
GN   T31E10.11 OR GA2OX3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Rounsley S.D., Kaul S., Lin X., Ketchum K.A., Crosby M.L.,
RA   Brandon R.C., Sykes S.M., Mason T.M., Kerlavage A.R., Adams M.D.,
RA   Somerville C.R., Venter J.C.;
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99218343; PubMed=10200325;
RA   Thomas S.G., Phillips A.L., Hedden P.;
RT   "Molecular cloning and functional expression of gibberellin 2-
RT   oxidases, multifunctional enzymes involved in gibberellin
RT   deactivation.";
RL   Proc. Natl. Acad. Sci. U.S.A. 96:4698-4703(1999).
DR   EMBL; AC004077; AAM14908.1; -.
DR   EMBL; AJ132437; CAB41009.1; -.
DR   InterPro; IPR005123; 2OG-FeII_Oxy.
DR   Pfam; PF03171; 2OG-FeII_Oxy; 1.
SQ   SEQUENCE   335 AA;  38216 MW;  181F6EAA1EE1C331 CRC64;

  Query Match             35.9%;  Score 51;  DB 10;  Length 335;
  Best Local Similarity   37.5%;  Pred. No. 37;
  Matches    9;  Conservative    9;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAQEYGRELRRMSDEFVDSFKKGL 27
              | :|| :|::||| :|::  :: |
Db        136 AVEEYIKEMKRMSSKFLEMVEEEL 159
```

```
RESULT 6
Q8TVX8
ID   Q8TVX8      PRELIMINARY;      PRT;   722 AA.
AC   Q8TVX8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Predicted ATPase of the AAA+ class.
GN   MK1260.
OS   Methanopyrus kandleri.
OC   Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC   Methanopyrus.
OX   NCBI_TaxID=2320;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AV19 / DSM 6324 / JCM 9639;
RX   MEDLINE=21927647; PubMed=11930014;
RA   Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA   Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA   Natale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stetter K.O.,
RA   Malykh A.G., Koonin E.V., Kozyavkin S.A.;
RT   "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT   and monophyly of archaeal methanogens.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
DR   EMBL; AE010416; AAM02473.1; -.
KW   Complete proteome.
SQ   SEQUENCE   722 AA;  81933 MW;  7FA609E9868A997E CRC64;

  Query Match             35.9%;  Score 51;  DB 17;  Length 722;
  Best Local Similarity   42.9%;  Pred. No. 88;
  Matches    9;  Conservative    7;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          5 AQEYGRELRRMSDEFVDSFKK 25
              |:|:|:|||| :|: : :|
Db        635 AREHGKELRRKEEEYRNRVRK 655


RESULT 7
Q9HZQ3
ID   Q9HZQ3      PRELIMINARY;      PRT;  1248 AA.
AC   Q9HZQ3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cobalamin biosynthetic protein CobN.
GN   COBN OR PA2944.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004720; AAG06332.1; -.
DR   InterPro; IPR003672; CobN/Mg_chltase.
DR   InterPro; IPR000408; Reg_chr_condens.
DR   Pfam; PF02514; cobN-Mg_chel; 1.
DR   PROSITE; PS00626; RCC1_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1248 AA;  138499 MW;  C3D3DBFEE6736C7A CRC64;

  Query Match             35.9%;  Score 51;  DB 16;  Length 1248;
  Best Local Similarity   47.8%;  Pred. No. 1.6e+02;
  Matches   11;  Conservative    5;  Mismatches    5;  Indels    2;  Gaps    1;

Qy          2 LWAAQEYG--RELRRMSDEFVDS 22
              |  |: ||  |:| |::||| |:
Db        611 LTRAESYGPLRDLERLADEFYDA 633


RESULT 8
Q8N955
ID   Q8N955      PRELIMINARY;      PRT;   173 AA.
AC   Q8N955;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein FLJ38343.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RA   Kanehori K., Ishibashi T., Chiba Y., Fujimori K., Hiraoka S.,
RA   Tanai H., Watanabe S., Ishida S., Ono Y., Hotuta T., Watanabe M.,
RA   Sugiyama T., Irie R., Otsuki T., Sato H., Ota T., Wakamatsu A.,
RA   Ishii S., Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA   Kimura K., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA   Wagatsuma M., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA   Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT   "NEDO human cDNA sequencing project.";
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AK095662; BAC04601.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   173 AA;  21285 MW;  0EBFF3BFD96FE775 CRC64;

  Query Match             35.2%;  Score 50;  DB 4;  Length 173;
  Best Local Similarity   47.6%;  Pred. No. 24;
  Matches   10;  Conservative    4;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          6 QEYGRELRRMSDEFVDSFKKG 26
              :| ||| ||  : ::|  |||
Db         36 KEGGREERRKEERWIDGRKKG 56


RESULT 9
Q8XXS6
ID   Q8XXS6      PRELIMINARY;      PRT;   230 AA.
AC   Q8XXS6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Probable ATP-binding ABC transporter protein.
GN   RSC2037 OR RS03602.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Ralstoniaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646067; CAD15739.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
```

```
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
KW   Complete proteome.
SQ   SEQUENCE   230 AA;  25231 MW;  7C3FDA1E7A19A2A4 CRC64;

  Query Match             35.2%;  Score 50;  DB 16;  Length 230;
  Best Local Similarity   71.4%;  Pred. No. 34;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 QEYGRELRRMSDEF 19
              || || |||: |||
Db        168 QEIGRTLRRLVDEF 181


RESULT 10
Q8F5R1
ID   Q8F5R1      PRELIMINARY;      PRT;   330 AA.
AC   Q8F5R1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   DTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
GN   RFBB3 OR LA1606.
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA   Ren S.;
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE011338; AAN48805.1; -.
KW   Lyase; Complete proteome.
SQ   SEQUENCE   330 AA;  36561 MW;  88B495C834E497C8 CRC64;

  Query Match             35.2%;  Score 50;  DB 16;  Length 330;
  Best Local Similarity   39.3%;  Pred. No. 51;
  Matches   11;  Conservative    6;  Mismatches    9;  Indels    2;  Gaps    1;

Qy          2 LWAAQEYGRELRRMSDEF--VDSFKKGL 27
              |||: :  ::|   | |:   | ||:||
Db        280 LWASNQKAKDLINWSPEYGGRDGFKRGL 307


RESULT 11
Q8EH69
ID   Q8EH69      PRELIMINARY;      PRT;   363 AA.
AC   Q8EH69;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyr-sensitive.
GN   AROF OR SO1361.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA   Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA   DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA   Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA   Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA   Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA   Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015580; AAN54426.1; -.
DR   TIGR; SO1361; -.
KW   Complete proteome.
SQ   SEQUENCE   363 AA;  39874 MW;  D59DE879EA9BFCB6 CRC64;

  Query Match             35.2%;  Score 50;  DB 16;  Length 363;
  Best Local Similarity   45.0%;  Pred. No. 56;
  Matches    9;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          4 AAQEYGRELRRMSDEFVDSF 23
              ||:||  :|::: ||  | |
Db         69 AAKEYALKLKKLHDELSDEF 88


RESULT 12
Q9U1L1
ID   Q9U1L1      PRELIMINARY;      PRT;   371 AA.
AC   Q9U1L1;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   EG:BACR7A4.8 protein.
GN   EG:BACR7A4.8 OR CG3021.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Papagiannakis G., Spanos L., Bolshakov V., Siden-Kiamos I., Louis C.;
RT   "Sequencing the distal X chromosome of Drosophila melanogaster.";
RL   Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Benos P.;
RL   Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL109630; CAB65878.1; -.
DR   FlyBase; FBgn0040337; EG:BACR7A4.8.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
DR   TIGRFAMs; TIGR00420; trmU; 1.
SQ   SEQUENCE   371 AA;  41315 MW;  759BA109E3C0AE36 CRC64;

  Query Match             35.2%;  Score 50;  DB 5;  Length 371;
  Best Local Similarity   28.6%;  Pred. No. 58;
  Matches   10;  Conservative    9;  Mismatches    6;  Indels   10;  Gaps    1;

Qy          3 WAAQEYGRELRRMS----------DEFVDSFKKGL 27
              || :: | |||:::           :|:| :: ||
Db         57 WACRQLGVELRQVNYVREYWTAVFSQFLDDYQMGL 91


RESULT 13
Q9W5B6
ID   Q9W5B6      PRELIMINARY;      PRT;   389 AA.
AC   Q9W5B6; Q960Y5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   CG3021 protein (LD29918p).
GN   EG:BACR7A4.8 OR CG3021.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
```

```
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA   Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA   Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA   Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA   Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA   Ferriera S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA   Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA   Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA   McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA   Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA   Phouanenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA   Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA   Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT   "Sequencing of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA   Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA   Tupy J.L., Bergman C., Berman B., Carlson J.W., Celniker S.E.,
RA   Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA   Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA   Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA   Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT   "Annotation of Drosophila melanogaster genome.";
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL   Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   FlyBase;
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE003419; AAF45578.3; -.
DR   EMBL; AY051773; AAK93197.1; -.
DR   FlyBase; FBgn0040337; EG:BACR7A4.8.
DR   InterPro; IPR004506; TrmU.
DR   Pfam; PF03054; tRNA_Me_trans; 1.
DR   TIGRFAMs; TIGR00420; trmU; 1.
SQ   SEQUENCE   389 AA;  43350 MW;  20C9405FFCC9FAE2 CRC64;

  Query Match             35.2%;  Score 50;  DB 5;  Length 389;
  Best Local Similarity   28.6%;  Pred. No. 61;
  Matches   10;  Conservative    9;  Mismatches    6;  Indels   10;  Gaps    1;

Qy          3 WAAQEYGRELRRMS----------DEFVDSFKKGL 27
              || :: | |||:::           :|:| :: ||
Db         57 WACRQLGVELRQVNYVREYWTAVFSQFLDDYQMGL 91


RESULT 14
Q95U84
ID   Q95U84      PRELIMINARY;      PRT;   320 AA.
AC   Q95U84;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   GH01229p.
GN   CG7993.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Berkeley;
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA   Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA   Nunoo J., Pacleb J., Paragas V., Park S., Phouanenavong S., Wan K.,
RA   Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL   Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY058248; AAL13477.1; -.
DR   FlyBase; FBgn0038585; CG7993.
DR   InterPro; IPR007109; Brix.
DR   Pfam; PF04427; Brix; 1.
DR   PROSITE; PS50833; BRIX; 1.
SQ   SEQUENCE   320 AA;  36509 MW;  EE98936DD68B3703 CRC64;

  Query Match             34.9%;  Score 49.5;  DB 5;  Length 320;
  Best Local Similarity   43.5%;  Pred. No. 58;
  Matches   10;  Conservative    6;  Mismatches    4;  Indels    3;  Gaps    1;

Qy          3 WAAQEYGRELRRMSDEFVDSFKK 25
              ||  |   ||||: : |:|:|::
Db        149 WAQTE---ELRRLRNLFIDTFQR 168


RESULT 15
Q8KDC3
ID   Q8KDC3      PRELIMINARY;      PRT;   214 AA.
AC   Q8KDC3;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein CT1131.
GN   CT1131.
OS   Chlorobium tepidum.
```

```
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobium.
OX   NCBI_TaxID=1097;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA   Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA   Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA   Vamathevan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA   Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR   EMBL; AE012874; AAM72364.1; -.
DR   TIGR; CT1131; -.
DR   InterPro; IPR002785; DUF83.
DR   TIGRFAMs; TIGR00372; TIGR00372; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   214 AA;  24004 MW;  56D389219D60B6AD CRC64;

  Query Match            34.5%;  Score 49;  DB 16;  Length 214;
  Best Local Similarity  37.0%;  Pred. No. 43;
  Matches   10;  Conservative    8;  Mismatches    7;  Indels    2;  Gaps    1;

Qy          1 NLWAAQEYGRELRRMSDEFVDSFKKGL 27
              ||: |:  |||:   :|  | |:::|:
Db         35 NLYTAE--GREMHERADSAVTSYREGV 59


Search completed: September 23, 2003, 09:45:06
Job time : 101 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 38.1857 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145	100.0	27	21	AAB37003	Bcl2 polypeptide B
2	145	100.0	162	22	AAB70370	Shorter murine BAD
3	145	100.0	204	17	AAR95168	bcl-x(L)/bcl-2 ass
4	145	100.0	204	19	AAW61315	Murine BCL-XL/BCL-
5	145	100.0	204	19	AAW61316	Mutant BCL-XL/BCL-
6	145	100.0	204	19	AAW61317	Mutant BCL-XL/BCL-
7	145	100.0	204	19	AAW61318	Mutant BCL-XL/BCL-
8	145	100.0	204	19	AAW58832	Murine BAD protein
9	145	100.0	204	22	AAB70369	Longer murine BAD
10	145	100.0	204	24	ABR39082	Murine BAD protein
11	145	100.0	567	22	AAU00220	Bad-DTTR apoptosis
12	138	95.2	26	21	AAB37001	Bcl2 polypeptide B
13	138	95.2	26	21	AAB37002	Bcl2 polypeptide B
14	138	95.2	27	21	AAB37056	Bcl2 polypeptide B
15	138	95.2	28	21	AAB37055	Bcl2 polypeptide B
16	120.5	83.1	166	18	AAW32476	BBC6 protein for r
17	120.5	83.1	168	19	AAW55779	Human Bcl-xL/Bcl-2
18	120.5	83.1	168	21	AAB13512	Human cell prolife
19	120.5	83.1	168	22	AAB70368	Human BAD mutant a
20	120.5	83.1	168	22	AAB48287	Human Bad protein.
21	120.5	83.1	168	22	AAG67688	Amino acid sequenc
22	120.5	83.1	168	24	ABR39081	Human BAD protein
23	120.5	83.1	201	23	ABP41630	Human ovarian anti
24	114	78.6	24	23	AAU78627	Human Bad peptide
25	114	78.6	25	23	ABP56161	PTPC-interacting T
26	114	78.6	25	23	ABG78484	Mutant Bcl2 compet
27	114	78.6	25	23	ABG78493	Mutant Bcl2 compet
28	114	78.6	25	23	AAU78610	Human Bad peptide
29	114	78.6	25	23	AAU78620	Human Bad peptide
30	113	77.9	23	17	AAR95166	bcl-x(L)/bcl-2 ass
31	111	76.6	25	23	ABG78490	Mutant Bcl2 compet
32	111	76.6	25	23	AAU78617	Human Bad peptide
33	110	75.9	25	23	ABG78488	Mutant Bcl2 compet
34	110	75.9	25	23	ABG78489	Mutant Bcl2 compet
35	110	75.9	25	23	AAU78615	Human Bad peptide
36	110	75.9	25	23	AAU78616	Human Bad peptide
37	109	75.2	23	23	AAU78628	Human Bad peptide
38	109	75.2	25	23	ABG78486	Mutant Bcl2 compet
39	109	75.2	25	23	ABG78492	Mutant Bcl2 compet
40	109	75.2	25	23	ABG78497	Mutant Bcl2 compet
41	109	75.2	25	23	AAU78612	Human Bad peptide
42	109	75.2	25	23	AAU78619	Human Bad peptide
43	109	75.2	25	23	AAU78624	Human Bad peptide
44	108	74.5	25	23	ABG78485	Mutant Bcl2 compet
45	108	74.5	25	23	AAU78611	Human Bad peptide

ALIGNMENTS

RESULT 1
AAB37003
ID AAB37003 standard; peptide; 27 AA.
XX
AC AAB37003;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #3.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.

```
XX
PT   New peptide conjugates for modulating apoptosis or for inhibiting B
PT   cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT   treating neurodegenerative disorders, stroke, or cancer -
XX
PS   Claim 18; Page 17; 74pp; English.
XX
CC   The invention relates to a peptide conjugate having the formula:
CC   (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC   to the N-terminus of the peptide, or a side chain of the peptide where
CC   the functional group of the side chain is NH2 or OH; or X = O or NH,
CC   when the R-X group is attached to the C-terminus of the peptide, or a
CC   side chain of the peptide, where the side chain functional group is COOH
CC   or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC   or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC   monosubstituted with a 1-5C straight or branched chain alkyl group,
CC   phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC   alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC   of the peptide portion of the conjugate. The peptides represent analogues
CC   of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC   the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC   useful for modulating apoptosis in the cells of a subject, or for
CC   reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC   apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC   function. In particular, the peptide conjugate is useful for treating a
CC   subject afflicted with a cancer characterized by cancer cells that
CC   express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC   non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC   acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC   conjugate is also useful for treating disorders characterized by
CC   increased apoptosis, e.g. neurodegenerative disorders, acquired
CC   immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ   Sequence   27 AA;

  Query Match             100.0%;  Score 145;  DB 21;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 2.6e-15;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27


RESULT 2
AAB70370
ID   AAB70370 standard; protein; 162 AA.
XX
AC   AAB70370;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Shorter murine BAD mutant amino acid sequence SEQ ID NO:3.
XX
KW   Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW   immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW   cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW   immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW   immunodeficiency disease; neurodegenerative disease; viral infection;
KW   ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW   lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS   Mus musculus.
OS   Synthetic.
XX
PN   WO200110888-A1.
XX
PD   15-FEB-2001.
XX
PF   30-MAY-2000; 2000WO-US11864.
XX
PR   28-MAY-1999;   99US-0136783.
XX
PA   (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI   Zhou X;
XX
DR   WPI; 2001-138734/14.
XX
PT   New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT   useful for screening for candidate compounds which induce or inhibit
PT   apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT   Ser113 -
XX
PS   Claim 7; Page 148-149; 157pp; English.
XX
CC   The present invention describes an isolated or synthetic polypeptide
CC   (I) comprising a less than full length amino acid sequence of a mutant
CC   Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC   fragment, which contains amino acid substitutions at Ser118 of a human
CC   BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC   BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC   nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC   antiarthritic, antiinflammatory and immunosuppressive activities, and
CC   can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC   polynucleotides can be used for screening candidate compounds and drugs
CC   for activity that promote cell survival or apoptosis. Other uses include
CC   inducing or inhibiting apoptosis in a cell. Candidate compounds
CC   identified and (mutant) BAD polypeptides are useful in treating
CC   immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC   death, reperfusion cell death, wound healing, cancer, viral infections,
CC   lymphoproliferative conditions, arthritis, infertility, inflammation and
CC   autoimmune diseases. The present sequence represents a specifically
CC   claimed shorter murine BAD mutant amino acid sequence from the present
CC   invention.
XX
SQ   Sequence   162 AA;

  Query Match             100.0%;  Score 145;  DB 22;  Length 162;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db         98 NLWAAQRYGRELRRMSDEFEGSFKGLP 124


RESULT 3
AAR95168
ID   AAR95168 standard; Protein; 204 AA.
XX
AC   AAR95168;
XX
DT   06-JAN-1997  (first entry)
XX
DE   bcl-x(L)/bcl-2 associated death promoter protein.
XX
KW   Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW   polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW   cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW   neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS   Mus musculus.
XX
FH   Key             Location/Qualifiers
FT   Region          147..149
FT                   /note= "BH1 conserved amino acids"
FT   Region          191..192
FT                   /note= "BH2 conserved amino acids"
FT   Domain          38..61
FT                   /note= "PEST sequence"
FT   Domain          111..130
FT                   /note= "PEST sequence"
XX
PN   WO9613614-A1.
XX
```

PD 09-MAY-1996.
XX
PF 31-OCT-1995; 95WO-US14246.
XX
PR 31-OCT-1994; 94US-0333565.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1996-251465/25.
DR N-PSDB; AAT29479.
XX
PT Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT useful to treat neoplasia and apoptosis and to identify agents
PT inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS Claim 3; Fig 1; 130pp; English.
XX
CC This sequence represents the murine bcl-x(L)/bcl-2 associated death
CC promoter (Bad) gene. Bad is a 22.1 kD protein which interacts with
CC bcl-2 and bcl-x proteins and regulates cell death. It has homology
CC to the bcl-2-related family clustered in the BH1 and BH2 domain. Bad
CC has been found to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid
CC assays and in vivo in mammalian cells. Overexpressed Bad counters the
CC death inhibitory activity of bcl-x(L), but is much less effective at
CC countering the death inhibitory activity of bcl-2. Bad expression can
CC accelerate apoptotic cell death induced by cytokine deprivation in an
CC IL-3 dependent cell line expressing bcl-x(L), and its also counters the
CC death repressor activity of bcl-x(L). Bad competes with Bax for binding
CC to bcl-x(L). Bad may be used to identify agents which inhibit its
CC binding to bcl-2 or bcl-x(L) to form heterodimers. Such agents may be
CC used to treat neurodegenerative diseases, immunodeficiency diseases,
CC e.g. AIDS, senescence or ischaemia.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 17; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 4
AAW61315
ID AAW61315 standard; Protein; 204 AA.
XX
AC AAW61315;
XX
DT 07-OCT-1998 (first entry)
XX
DE Murine BCL-XL/BCL-2 associated cell death regulator.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27833.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 1; Fig 10; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence is the murine BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 5
AAW61316
ID AAW61316 standard; Protein; 204 AA.
XX
AC AAW61316;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #1.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.

DR N-PSDB; AAV27834.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 59; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 6
AAW61317
ID AAW61317 standard; Protein; 204 AA.
XX
AC AAW61317;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #2.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27835.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 60; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 7
AAW61318
ID AAW61318 standard; Protein; 204 AA.
XX
AC AAW61318;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #3.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27836.
XX

PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 60-61; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 8
AAW58832
ID AAW58832 standard; protein; 204 AA.
XX
AC AAW58832;
XX
DT 23-JUL-1998 (first entry)
XX
DE Murine BAD protein.
XX
KW BAD protein; Bcl-XL/Bcl-2 associated cell death regulator; 14-3-3;
KW serine phosphorylation; post-translational modification; apoptosis;
KW signal transduction regulator; phosphoserine phosphatase; senescence;
KW immunodeficiency disease; neurodegenerative disease; infertility;
KW cancer, viral infection; lymphoproliferative condition; arthritis;
KW inflammation; autoimmune diseases.
XX
OS Mus sp.
XX
PN WO9809643-A1.
XX
PD 12-MAR-1998.
XX
PF 09-SEP-1997; 97WO-US15871.
XX
PR 09-SEP-1996; 96US-0707868.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-207049/18.
XX
PT Serine-phosphorylated Bcl-X-l/Bcl-2 Associated cell Death regulator
PT polypeptide - useful for modulation of apoptosis associated with,
PT e.g. cancer and immunodeficiency diseases
XX
PS Claim 3; Fig 8; 61pp; English.
XX
CC This sequence represents a novel serine-phosphorylated protein, BAD
CC (Bcl-Xl/Bcl-2 associated cell death regulator). The serine residue is
CC phosporylated in a post-translational modification and allows binding
CC to the 14-3-3 protein which is a signal transduction regulator.
CC Modulators of phosphorylated BAD, which act through inhibition/activation
CC of a phosphoserine phosphatase, are useful for preventing/treating
CC increased/decreased apoptosis in a cell. The increased apoptosis may
CC result from immunodeficiency diseases, senescence, neurodegenerative
CC disease, ischaemic cell death, reperfusion cell death, infertility and
CC wound-healing. Decreased apoptosis may result from cancer, viral
CC infection, lymphoproliferative conditions, arthritis, infertility,
CC inflammation and autoimmune diseases. Measuring the amount of
CC phosphorylated compared to unphosphorylated BAD polypeptide and/or total
CC BAD in a cell is useful for determining the apoptotic state of a cell.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 9
AAB70369
ID AAB70369 standard; protein; 204 AA.
XX
AC AAB70369;
XX
DT 02-MAY-2001 (first entry)
XX
DE Longer murine BAD mutant amino acid sequence SEQ ID NO:2.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX

PS Claim 4; Page 148; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed longer murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 22; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 10
ABR39082
ID ABR39082 standard; Protein; 204 AA.
XX
AC ABR39082;
XX
DT 10-MAY-2003 (first entry)
XX
DE Murine BAD protein SEQ ID NO:4.
XX
KW Murine; BAD; herpes simplex virus; HSV; US3; herpes virus; apoptosis;
KW virucide; infection.
XX
OS Mus musculus.
XX
PN WO2003012049-A2.
XX
PD 13-FEB-2003.
XX
PF 31-JUL-2002; 2002WO-US24177.
XX
PR 31-JUL-2001; 2001US-308929P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Munger J, Roizman B;
XX
DR WPI; 2003-248168/24.
DR N-PSDB; ABZ81201.
XX
PT Inducing apoptosis in a cell infected with herpes simplex virus, HSV,
PT by administering to the cell, a composition comprising an agent that
PT inhibits phosphorylation of pro-apoptotic polypeptide BAD by HSV US3 -
XX
PS Claim 15; Page 168; 192pp; English.
XX
CC The present invention describes a method (M1) for inducing apoptosis in
CC a cell infected with herpes simplex virus (HSV), which comprises
CC administering to the cell, a composition having an agent that inhibits
CC phosphorylation of pro-apoptotic polypeptide BAD by HSV US3. Also
CC described is a method (M2) for treating a patient infected with HSV, by
CC administering to the patient, a composition comprising a peptide
CC comprising a sequence of 4-100 continuous amino acids of a 168 residue
CC amino acid sequence (see ABR39081), where the peptide comprises ser112,
CC ser135, or ser155, or their combinations. BAD has virucide activity.
CC M1 is useful for inducing apoptosis in a cell infected with HSV, where
CC the cell is in a human. M2 is useful for treating a patient infected
CC with HSV. The present sequence represents murine BAD, which is used in
CC the exemplification of the present invention.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 145; DB 24; Length 204;
Best Local Similarity 100.0%; Pred. No. 2.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166

RESULT 11
AAU00220
ID AAU00220 standard; Protein; 567 AA.
XX
AC AAU00220;
XX
DT 31-MAY-2001 (first entry)
XX
DE Bad-DTTR apoptosis-modifying fusion protein.
XX
KW Mouse; Bad-DTTR; apoptosis; cancer; spinal muscular atrophy;
KW diphtheria toxin receptor binding domain; DTR; neoplasm; tumour;
KW hyper-proliferation; Alzheimer's disease; neurodegenerative disorder;
KW transient ischaemic neuronal injury; stroke; spinal cord injury;
KW Huntington's disease.
XX
OS Chimeric - Mus sp.
OS Chimeric - Corynebacterium diptheriae.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Region 3..12
FT /note= "10x histidine tag"
XX
PN WO200112661-A2.
XX
PD 22-FEB-2001.
XX
PF 15-AUG-2000; 2000WO-US22293.
XX
PR 16-AUG-1999; 99US-0149220.
XX
PA (HARD) HARVARD COLLEGE.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Youle RJ, Liu X, Collier RJ;
XX
DR WPI; 2001-218343/22.
DR N-PSDB; AAS00248.
XX
PT Novel fusion protein for modifying apoptosis in target cell and
PT reducing apoptosis after transient ischaemic neuronal injury, has two
PT domains which targets protein to a cell and modifies apoptotic response
PT of cell -
XX
PS Claim 4; Page 59-61; 65pp; English.
XX
CC The sequence represents the amino acid sequence of Bad-DTTR apoptosis-
CC modifying fusion protein comprising Bad gene sequence fused via a short
CC linker to diphtheria toxin translocation domain (DTTR). The
CC functional apoptosis-modifying fusion protein is capable of binding a
CC target cell and integrating into or crossing a cellular membrane of the

CC target cell. The apoptosis-modifying fusion protein comprises at least
CC two domains: the DTR domain, which targets the fusion protein to the
CC target cell and the Bcl-XL domain, which modifies an apoptotic response
CC of the target cell. The fusion protein is useful for modifying
CC (inhibiting or enhancing) apoptosis in a target cell, such as neuron,
CC lymphocyte, cancer, neoplasm, macrophage, epithelial, stem, tumour or
CC hyper-proliferative cell or an adipocyte. It is also useful for reducing
CC apoptosis in a subject after transient ischaemic neuronal injury,
CC especially spinal cord injury. The fusion protein may be used to treat
CC various diseases and injury conditions through inhibition or enhancement
CC of apoptotic cellular response, including neurodegenerative disorders
CC such as Alzheimer's disease, Huntington's disease, spinal muscular
CC atrophy, stroke episodes and unregulated cell growth as in tumours and
CC various cancers. The apoptosis-modifying fusion protein can be delivered
CC effectively throughout the body and targeted to selective tissue and
CC cells.
XX
SQ Sequence 567 AA;

Query Match 100.0%; Score 145; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 7.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
||||||||||||||||||||||||||||
Db 161 NLWAAQRYGRELRRMSDEFEGSFKGLP 187

RESULT 12
AAB37001
ID AAB37001 standard; peptide; 26 AA.
XX
AC AAB37001;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #1.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 95.2%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 13
AAB37002
ID AAB37002 standard; peptide; 26 AA.
XX
AC AAB37002;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #2.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where

CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 95.2%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 14
AAB37056
ID AAB37056 standard; peptide; 27 AA.
XX
AC AAB37056;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #56.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX

CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 95.2%; Score 138; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 3.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
||||||||||||||||||||||||||
Db 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 15
AAB37055
ID AAB37055 standard; peptide; 28 AA.
XX
AC AAB37055;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #55.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -

XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 28 AA;

Query Match 95.2%; Score 138; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 3.2e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGL 26
              ||||||||||||||||||||||||||
Db          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
```

Search completed: September 15, 2003, 17:22:13
Job time : 38.1857 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 14.0786 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145	100.0	204	1	US-08-333-565-2	Sequence 2, Appli
2	145	100.0	204	2	US-08-661-479-2	Sequence 2, Appli
3	145	100.0	204	2	US-08-733-505A-1	Sequence 1, Appli
4	145	100.0	204	2	US-08-733-505A-12	Sequence 12, Appl
5	145	100.0	204	2	US-08-733-505A-13	Sequence 13, Appl
6	145	100.0	204	2	US-08-733-505A-14	Sequence 14, Appl
7	142	97.9	204	2	US-08-717-123-3	Sequence 3, Appli
8	142	97.9	204	4	US-09-375-257-3	Sequence 3, Appli
9	120.5	83.1	166	1	US-08-665-617-2	Sequence 2, Appli
10	120.5	83.1	168	2	US-08-717-123-2	Sequence 2, Appli
11	120.5	83.1	168	3	US-08-985-335-1	Sequence 1, Appli
12	120.5	83.1	168	3	US-08-985-335-7	Sequence 7, Appli
13	120.5	83.1	168	3	US-09-410-372-1	Sequence 1, Appli
14	120.5	83.1	168	3	US-09-410-372-7	Sequence 7, Appli
15	120.5	83.1	168	4	US-09-375-257-2	Sequence 2, Appli
16	113	77.9	23	1	US-08-333-565-10	Sequence 10, Appl
17	113	77.9	23	2	US-08-661-479-10	Sequence 10, Appl
18	102	70.3	59	2	US-08-733-505A-55	Sequence 55, Appl
19	102	70.3	59	2	US-08-733-505A-56	Sequence 56, Appl
20	102	70.3	59	2	US-08-733-505A-57	Sequence 57, Appl
21	102	70.3	59	2	US-08-733-505A-58	Sequence 58, Appl
22	86	59.3	16	1	US-08-333-565-26	Sequence 26, Appl
23	86	59.3	16	2	US-08-661-479-26	Sequence 26, Appl
24	61	42.1	11	2	US-08-733-505A-34	Sequence 34, Appl
25	61	42.1	11	2	US-08-706-741B-69	Sequence 69, Appl
26	61	42.1	11	2	US-08-924-695A-69	Sequence 69, Appl
27	53	36.6	946	3	US-09-074-579-3	Sequence 3, Appli
28	53	36.6	946	3	US-09-388-774-3	Sequence 3, Appli
29	53	36.6	946	4	US-09-546-153-1	Sequence 1, Appli
30	51	35.2	66	2	US-08-867-087B-40	Sequence 40, Appl
31	50.5	34.8	467	4	US-09-252-991A-18296	Sequence 18296, A
32	49	33.8	263	4	US-09-651-656-27	Sequence 27, Appl
33	49	33.8	263	4	US-09-650-855-27	Sequence 27, Appl
34	48.5	33.4	904	4	US-09-328-352-4656	Sequence 4656, Ap
35	46.5	32.1	415	4	US-09-252-991A-31348	Sequence 31348, A
36	46	31.7	322	4	US-09-252-991A-30979	Sequence 30979, A
37	46	31.7	610	4	US-09-252-991A-19594	Sequence 19594, A
38	45.5	31.4	351	4	US-09-252-991A-32466	Sequence 32466, A
39	45.5	31.4	906	4	US-09-252-991A-31458	Sequence 31458, A
40	45	31.0	229	4	US-09-252-991A-23807	Sequence 23807, A
41	45	31.0	303	4	US-09-328-352-5164	Sequence 5164, Ap
42	45	31.0	356	4	US-09-235-103-2	Sequence 2, Appli
43	45	31.0	356	4	US-09-235-103-4	Sequence 4, Appli
44	45	31.0	513	1	US-08-464-340A-2	Sequence 2, Appli
45	45	31.0	513	5	PCT-US94-08449A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
; GENERAL INFORMATION:
; APPLICANT: KORSMEYER, Stanley J.
; TITLE OF INVENTION: Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
; TITLE OF INVENTION: REGULATOR
; NUMBER OF SEQUENCES: 59
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Khourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,565
; FILING DATE: 31-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 15726A-000700
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 204 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..204
; OTHER INFORMATION: /note= "Deduced amino acid sequence
; OTHER INFORMATION: of mouse BAD."
US-08-333-565-2

Query Match 100.0%; Score 145; DB 1; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.4e-14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 2
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             100.0%;  Score 145;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 3
US-08-733-505A-1
; Sequence 1, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-733-505A-1

  Query Match             100.0%;  Score 145;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 4
US-08-733-505A-12
; Sequence 12, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-12

  Query Match             100.0%;  Score 145;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 5
US-08-733-505A-13
; Sequence 13, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-13

  Query Match             100.0%;  Score 145;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 6
US-08-733-505A-14
; Sequence 14, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-14

  Query Match             100.0%;  Score 145;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 7
US-08-717-123-3
; Sequence 3, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
```

```
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-717-123-3

  Query Match             97.9%;  Score 142;  DB 2;  Length 204;
  Best Local Similarity   96.3%;  Pred. No. 4.1e-14;
  Matches   26;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGLP 166


RESULT 8
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match             97.9%;  Score 142;  DB 4;  Length 204;
  Best Local Similarity   96.3%;  Pred. No. 4.1e-14;
  Matches   26;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGLP 166


RESULT 9
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match             83.1%;  Score 120.5;  DB 1;  Length 166;
  Best Local Similarity   89.3%;  Pred. No. 6.2e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        101 NLWAAQRYGRELRRMSDEFVDSFKKGLP 128


RESULT 10
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match             83.1%;  Score 120.5;  DB 2;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130
```

```
RESULT 11
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match           83.1%;  Score 120.5;  DB 3;  Length 168;
  Best Local Similarity  89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 12
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match           83.1%;  Score 120.5;  DB 3;  Length 168;
  Best Local Similarity  89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 13
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match             83.1%;  Score 120.5;  DB 3;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||  |
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 14
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match             83.1%;  Score 120.5;  DB 3;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 15
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match             83.1%;  Score 120.5;  DB 4;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 6.3e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


Search completed: September 15, 2003, 17:45:05
Job time : 14.0786 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 21.4071 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	142	97.9	204	9	US-09-922-378-3	Sequence 3, Appli
2	142	97.9	204	14	US-10-066-179-3	Sequence 3, Appli
3	120.5	83.1	168	9	US-09-922-378-2	Sequence 2, Appli
4	120.5	83.1	168	9	US-09-894-657-1	Sequence 1, Appli
5	120.5	83.1	168	9	US-09-894-657-7	Sequence 7, Appli
6	120.5	83.1	168	14	US-10-066-179-2	Sequence 2, Appli
7	114	78.6	25	15	US-10-059-261-258	Sequence 258, App
8	71	49.0	15	15	US-10-174-105A-147	Sequence 147, App
9	53	36.6	946	9	US-09-828-423-3	Sequence 3, Appli
10	50	34.5	215	15	US-10-156-761-9145	Sequence 9145, Ap
11	47	32.4	35	15	US-10-092-750-1	Sequence 1, Appli
12	47	32.4	138	15	US-10-092-750-241	Sequence 241, App
13	46	31.7	682	12	US-10-238-075-1077	Sequence 1077, Ap
14	45.5	31.4	852	9	US-09-752-639-153	Sequence 153, App
15	45.5	31.4	852	10	US-09-984-198-153	Sequence 153, App
16	45	31.0	513	12	US-10-199-869-4	Sequence 4, Appli
17	45	31.0	513	14	US-10-143-002-2	Sequence 2, Appli
18	45	31.0	513	15	US-10-325-891-2	Sequence 2, Appli
19	44.5	30.7	144	12	US-09-903-190-160	Sequence 160, App
20	44.5	30.7	146	15	US-10-121-757B-12	Sequence 12, Appl
21	44.5	30.7	160	10	US-09-738-973-435	Sequence 435, App
22	44.5	30.7	160	10	US-09-854-133-435	Sequence 435, App
23	44.5	30.7	160	15	US-10-144-649A-435	Sequence 435, App
24	44.5	30.7	309	15	US-10-102-806-649	Sequence 649, App
25	44.5	30.7	334	10	US-09-794-715A-8	Sequence 8, Appli
26	44.5	30.7	334	12	US-10-286-581-8	Sequence 8, Appli
27	44.5	30.7	334	15	US-10-046-924-8	Sequence 8, Appli
28	44	30.3	81	9	US-09-925-297-560	Sequence 560, App
29	44	30.3	272	15	US-10-156-761-11541	Sequence 11541, A
30	44	30.3	426	9	US-09-815-242-5704	Sequence 5704, Ap
31	44	30.3	699	14	US-10-008-355-8	Sequence 8, Appli
32	44	30.3	705	9	US-09-815-242-12463	Sequence 12463, A
33	44	30.3	712	14	US-10-008-355-2	Sequence 2, Appli
34	44	30.3	877	12	US-10-369-294-20	Sequence 20, Appl
35	43.5	30.0	543	15	US-10-156-761-13485	Sequence 13485, A
36	43	29.7	213	9	US-09-843-846-18	Sequence 18, Appl
37	43	29.7	232	10	US-09-881-752A-238	Sequence 238, App
38	43	29.7	270	11	US-09-934-455-162	Sequence 162, App
39	43	29.7	380	9	US-09-149-045-2	Sequence 2, Appli
40	43	29.7	380	15	US-10-166-359-2	Sequence 2, Appli
41	43	29.7	380	15	US-10-166-113-2	Sequence 2, Appli
42	43	29.7	380	15	US-10-166-357-2	Sequence 2, Appli
43	43	29.7	380	15	US-10-166-372-2	Sequence 2, Appli
44	43	29.7	380	15	US-10-184-722-3	Sequence 3, Appli
45	43	29.7	380	15	US-10-251-385-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE
; FILE REFERENCE: 480140.428D3
; CURRENT APPLICATION NUMBER: US/09/922,378
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-922-378-3

Query Match 97.9%; Score 142; DB 9; Length 204;
Best Local Similarity 96.3%; Pred. No. 2.5e-13;
Matches 26; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ||||||||||||||:||||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGLP 166
```

RESULT 2
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE

```
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             97.9%;  Score 142;  DB 14;  Length 204;
  Best Local Similarity   96.3%;  Pred. No. 2.5e-13;
  Matches   26;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||:||||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGLP 166


RESULT 3
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match             83.1%;  Score 120.5;  DB 9;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 3.2e-10;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 4
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Yue, Henry
;                   Lal, Preeti
;                   Shah, Purvi
;                   Corley, Neil C.
;        TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                            PROLIFERATION
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/894,657
;             FILING DATE: 28-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/410,372
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0421 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 168 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: SYNORAB01
;             CLONE: 358673
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             83.1%;  Score 120.5;  DB 9;  Length 168;
  Best Local Similarity   89.3%;  Pred. No. 3.2e-10;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 5
US-09-894-657-7
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Yue, Henry
;                   Lal, Preeti
;                   Shah, Purvi
;                   Corley, Neil C.
;        TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                            PROLIFERATION
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/894,657
;             FILING DATE: 28-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/410,372
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0421 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
```

```
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 168 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GenBank
;             CLONE: 1683637
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match            83.1%;  Score 120.5;  DB 9;  Length 168;
  Best Local Similarity  89.3%;  Pred. No. 3.2e-10;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 6
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match            83.1%;  Score 120.5;  DB 14;  Length 168;
  Best Local Similarity  89.3%;  Pred. No. 3.2e-10;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 7
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match            78.6%;  Score 114;  DB 15;  Length 25;
  Best Local Similarity  91.7%;  Pred. No. 4.1e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFK 24
              ||||||||||||||||||||  |||
Db          1 NLWAAQRYGRELRRMSDEFVDSFK 24


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECI
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHI
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match            49.0%;  Score 71;  DB 15;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.00058;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFEGS 22
              ||||||||||||||
Db          1 GRELRRMSDEFEGS 14


RESULT 9
US-09-828-423-3
; Sequence 3, Application US/09828423
; Patent No. US20020099178A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Guegler, Karl J.
;                   Patterson, Chandra
;        TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
;                            INHIBITOR HEAVY CHAIN PRECURSOR
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
```

```
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/828,423
;             FILING DATE: 05-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/388,774
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER: PF-0505 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;             TELEX: <Unknown>
;    INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 946 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GENEBANK
;             CLONE: gi33985
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-09-828-423-3

  Query Match             36.6%;  Score 53;  DB 9;  Length 946;
  Best Local Similarity   33.3%;  Pred. No. 20;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ::|  :  |     : | ||| | |:|
Db        212 DVWVIEPQGLRFLHVPDTFEGHFDGVP 238


RESULT 10
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

  Query Match             34.5%;  Score 50;  DB 15;  Length 215;
  Best Local Similarity   56.2%;  Pred. No. 12;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          6 QRYGRELRRMSDEFEG 21
              :|:| :|||| || :|
Db        108 ERWGGDLRRMRDEADG 123


RESULT 11
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-1

  Query Match             32.4%;  Score 47;  DB 15;  Length 35;
  Best Local Similarity   45.5%;  Pred. No. 5.1;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              :| ||    ||||: |||   :
Db         15 IWIAQ----ELRRIGDEFNAYY 32


RESULT 12
US-10-092-750-241
; Sequence 241, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-241

  Query Match             32.4%;  Score 47;  DB 15;  Length 138;
  Best Local Similarity   45.5%;  Pred. No. 21;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              :| ||    ||||: |||   :
Db         86 IWIAQ----ELRRIGDEFNAYY 103


RESULT 13
US-10-238-075-1077
; Sequence 1077, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
;  APPLICANT: I.N.S.E.R.M.
;  TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are is
;  TITLE OF INVENTION:  E.coli, and biological uses of these polynucleotides and of t
```

```
;  FILE REFERENCE: BLANDINE
;  CURRENT APPLICATION NUMBER: US/10/238,075
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 0003145
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 1576
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1077
;   LENGTH: 682
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-238-075-1077

  Query Match            31.7%;  Score 46;  DB 12;  Length 682;
  Best Local Similarity  40.0%;  Pred. No. 1.6e+02;
  Matches   10;  Conservative    5;  Mismatches   10;  Indels    0;  Gaps    0;

Qy          2 LWAAQRYGRELRRMSDEFEGSFKGL 26
              :|||||  | :: |: : |     ||
Db        610 IWAAQRNGAKVPRVRNGFTSMDIGL 634


RESULT 14
US-09-752-639-153
; Sequence 153, Application US/09752639
; Patent No. US20020091243A1
;  GENERAL INFORMATION:
;    APPLICANT:  Gatanaga, T.
;    APPLICANT:  Granger, G.A.
;    TITLE OF INVENTION:  Factors Altering Tumor Necrosis
;    TITLE OF INVENTION:  Factor Receptor Releasing Enzyme Activity, and Methods
;    TITLE OF INVENTION:  of Use Thereof
;    NUMBER OF SEQUENCES:  154
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/752,639
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US99/10793
;      FILING DATE:
;      APPLICATION NUMBER:  09/081,385
;      FILING DATE:
;      APPLICATION NUMBER:  08/964,747
;      FILING DATE:  05-NOV-1997
;      APPLICATION NUMBER:  60/030,761
;      FILING DATE:  06-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wu, Frank
;      REGISTRATION NUMBER:  41,386
;      REFERENCE/DOCKET NUMBER:  22000-20577.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  153:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  852 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-752-639-153

  Query Match            31.4%;  Score 45.5;  DB 9;  Length 852;
  Best Local Similarity  35.3%;  Pred. No. 2.4e+02;
  Matches   12;  Conservative    5;  Mismatches   10;  Indels    7;  Gaps    2;

Qy          1 NLWAA------QRY-GRELRRMSDEFEGSFKGLP 27
              ::|:        || ||:| |  | || :  | |
Db        552 DIWSTYLTKFIARYGGRKLERARDLFEQALDGCP 585


RESULT 15
US-09-984-198-153
; Sequence 153, Application US/09984198
; Patent No. US20020106679A1
;  GENERAL INFORMATION:
;    APPLICANT:  Gatanaga, T.
;    APPLICANT:  Granger, G.A.
;    TITLE OF INVENTION:  Factors Altering Tumor Necrosis
;    TITLE OF INVENTION:  Factor Receptor Releasing Enzyme Activity, and Methods
;    TITLE OF INVENTION:  of Use Thereof
;    NUMBER OF SEQUENCES:  154
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORRISON & FOERSTER
;      STREET:  755 PAGE MILL ROAD
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304-1018
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  Windows
;      SOFTWARE:  FastSEQ for Windows Version 2.0b
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/984,198
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US99/10793
;      FILING DATE:
;      APPLICATION NUMBER:  09/081,385
;      FILING DATE:
;      APPLICATION NUMBER:  08/964,747
;      FILING DATE:  05-NOV-1997
;      APPLICATION NUMBER:  60/030,761
;      FILING DATE:  06-NOV-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Wu, Frank
;      REGISTRATION NUMBER:  41,386
;      REFERENCE/DOCKET NUMBER:  22000-20577.21
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-813-5600
;      TELEFAX:  650-494-0792
;      TELEX:  706141
;  INFORMATION FOR SEQ ID NO:  153:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  852 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-984-198-153

  Query Match            31.4%;  Score 45.5;  DB 10;  Length 852;
  Best Local Similarity  35.3%;  Pred. No. 2.4e+02;
  Matches   12;  Conservative    5;  Mismatches   10;  Indels    7;  Gaps    2;

Qy          1 NLWAA------QRY-GRELRRMSDEFEGSFKGLP 27
              ::|:        || ||:| |  | || :  | |
Db        552 DIWSTYLTKFIARYGGRKLERARDLFEQALDGCP 585


Search completed: September 15, 2003, 17:47:52
Job time : 21.4071 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 12.15 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145	100.0	204	2	A55671	bad protein - mous
2	61	42.1	946	2	JC5575	inter-alpha-trypsi
3	60	41.4	946	2	S54354	inter-alpha-inhibi
4	53	36.6	223	2	D70760	hypothetical prote
5	53	36.6	946	1	IYHU2	inter-alpha-trypsi
6	52	35.9	370	2	S38185	2-dehydro-3-deoxy-
7	51	35.2	232	2	A42095	floral homeotic pr
8	50	34.5	374	2	C84338	spermidine/putresc
9	50	34.5	516	2	A96753	probable threonine
10	49	33.8	263	2	A64807	endonuclease VIII
11	49	33.8	263	2	A85572	hypothetical prote
12	49	33.8	263	2	C90721	hypothetical prote
13	49	33.8	263	2	AD0590	endonuclease VIII,
14	49	33.8	453	2	E83517	conserved hypothet
15	48.5	33.4	134	2	S40376	Ig kappa chain - h
16	48.5	33.4	314	2	T02975	annexin P35 - maiz
17	48	33.1	206	2	C36365	transforming prote
18	48	33.1	220	2	F72289	oxidoreductase, so
19	48	33.1	526	2	T08545	threonine synthase
20	47.5	32.8	779	2	B81287	hypothetical prote
21	47	32.4	597	2	G82308	oxaloacetate decar
22	47	32.4	967	2	F82668	oxoglutarate dehyd
23	47	32.4	5138	2	B96695	hypothetical prote
24	46.5	32.1	314	2	T02961	annexin P33 - maiz
25	46.5	32.1	435	2	A44308	Antho-RFamide prec
26	46.5	32.1	1140	2	T09486	hypothetical prote
27	46	31.7	399	2	T35440	probable polyamine
28	46	31.7	497	2	C98318	alkaline proteinas
29	46	31.7	1164	2	T24806	hypothetical prote
30	46	31.7	1378	2	A81393	DNA-directed RNA p
31	45.5	31.4	194	2	C75420	hypothetical prote
32	45.5	31.4	261	2	G69510	conserved hypothet
33	45.5	31.4	263	2	E81148	probable transfera
34	45.5	31.4	287	2	S43852	neuropeptide Pol-R
35	45.5	31.4	328	2	E83337	probable transmemb
36	45.5	31.4	334	2	A39172	Antho-RFamide neur
37	45.5	31.4	562	2	C71473	hypothetical prote
38	45.5	31.4	905	2	G83314	NADH dehydrogenase
39	45	31.0	165	2	S59899	chlorocruorin chai
40	45	31.0	295	2	F83201	conserved hypothet
41	45	31.0	346	2	H95406	conserved hypothet
42	45	31.0	513	2	JC5920	potassium channel
43	45	31.0	591	2	B44465	sodium ion pump ox
44	45	31.0	591	2	AB0509	oxaloacetate decar
45	45	31.0	591	2	AE0909	oxaloacetate decar

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 100.0%; Score 145; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 1e-13;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ||||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166
```

RESULT 2
JC5575
inter-alpha-trypsin inhibitor heavy chain 2 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
C;Accession: JC5575; PC4485
R;Nakatani, T.; Suzuki, Y.; Yamamoto, T.; Sinohara, H.
J. Biochem. 122, 71-82, 1997
A;Title: Molecular cloning and sequencing of cDNAs encoding three heavy-chain precurs
sin inhibitor heavy chain family.
A;Reference number: JC5574; MUID:97420688; PMID:9276673
A;Accession: JC5575
A;Molecule type: mRNA
A;Residues: 1-946 <NAK>
A;Cross-references: DDBJ:D89286; NID:g1694689; PIDN:BAA13939.1; PID:g1694690
A;Experimental source: liver
A;Accession: PC4485
A;Molecule type: protein
A;Residues: 55-64;140-146;151-156;424-447;500-528;577-605 <NA2>
C;Comment: In the plasma three inter-alpha-trypsin inhibitor heavy chains 1, 2 and 3
that the complexes play important role for panceatic cancer.
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
F;261-264,717-916/Disulfide bonds: #status predicted

Query Match 42.1%; Score 61; DB 2; Length 946;
Best Local Similarity 37.0%; Pred. No. 1;

Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |:|  :  |     : | ||| |:|:|
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGVP 238
```

RESULT 3
S54354
inter-alpha-inhibitor H2 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999
C;Accession: S54354
R;Chan, P.; Risler, J.L.; Raguenez, G.; Salier, J.P.
Biochem. J. 306, 505-512, 1995
A;Title: The three heavy-chain precursors for the inter-alpha-inhibitor family in mouse:
A;Reference number: S54353; MUID:95194326; PMID:7534067
A;Accession: S54354
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-946 <CHA>
A;Cross-references: EMBL:X70392; NID:g695633; PIDN:CAA49842.1; PID:g695634
C;Superfamily: inter-alpha-trypsin inhibitor complex component II

Query Match 41.4%; Score 60; DB 2; Length 946;
Best Local Similarity 37.0%; Pred. No. 1.4;
Matches 10; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |:|  :  |     : | ||| |:|:|
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGVP 238
```

RESULT 4
D70760
hypothetical protein Rv2014 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70760
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70760
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <COL>
A;Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CAA98415.1; PID:e1299911;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2014

Query Match 36.6%; Score 53; DB 2; Length 223;
Best Local Similarity 58.8%; Pred. No. 3.4;
Matches 10; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

```
Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181
```

RESULT 5
IYHU2
inter-alpha-trypsin inhibitor heavy chain 2 precursor - human
N;Alternate names: inter-alpha-trypsin inhibitor complex component II
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1990 #sequence_revision 30-Sep-1990 #text_change 04-Feb-2000
C;Accession: S00346; S09064; B39967; S00632; B34245; S28929; C53642
R;Gebhard, W.; Schreitmueller, T.; Hochstrasser, K.; Wachter, E.
FEBS Lett. 229, 63-67, 1988
A;Title: Complementary DNA and derived amino acid sequence of the precursor of one of
A;Reference number: S00346; MUID:88152237; PMID:2450046
A;Accession: S00346
A;Molecule type: mRNA
A;Residues: 1-946 <GEB>
A;Cross-references: EMBL:X07173
A;Experimental source: liver
A;Note: part of this sequence, including the amino end of the mature protein, was con
ssing occurs at the carboxyl as well as the amino end to produce the mature protein
A;Note: due to a double frameshift, the nucleic acid sequence of codons 363-372 is in
A;Note: in one clone, a T is lacking from codon 716; this clone could code for a prot
R;Schreitmueller, T.; Hochstrasser, K.; Reisinger, P.W.M.; Wachter, E.; Gebhard, W.
Biol. Chem. Hoppe-Seyler 368, 963-970, 1987
A;Title: cDNA cloning of human inter-alpha-trypsin inhibitor discloses three differen
A;Reference number: S09064; MUID:88024442; PMID:3663330
A;Accession: S09064
A;Molecule type: mRNA
A;Residues: 265,'RR',268-284,'D',286-946 <SCH>
A;Note: this sequence has been revised in reference S00346
R;Salier, J.P.; Diarra-Mehrpour, M.; Sesboue, R.; Bourguignon, J.; Benarous, R.; Ohku
Proc. Natl. Acad. Sci. U.S.A. 84, 8272-8276, 1987
A;Title: Isolation and characterization of cDNAs encoding the heavy chain of human in
A;Reference number: A39967; MUID:88068576; PMID:2446322
A;Accession: B39967
A;Molecule type: mRNA
A;Residues: 384-673,'A',675-704,'S',706-728,'D',730,'A',732-865 <SAL>
A;Cross-references: GB:M18193; GB:J03013; NID:g338222; PIDN:AAA60558.1; PID:g553647
A;Experimental source: liver
R;Salier, J.P.; Diarra-Mehrpour, M.; Sesbouee, R.; Bourguignon, J.; Martin, J.P.
Biol. Chem. Hoppe-Seyler 369(Suppl.), 15-18, 1988
A;Title: Human inter-alpha-trypsin inhibitor. Isolation and characterization of heavy
A;Reference number: S00632; MUID:89076497; PMID:2462430
A;Accession: S00632
A;Molecule type: mRNA
A;Residues: 384-673,'A',675-704,'S',706-728,'D',730,'A',732-766 <SA2>
A;Cross-references: GB:M33033; NID:g186589; PIDN:AAA59195.1; PID:g186590
R;Enghild, J.J.; Thogersen, I.B.; Pizzo, S.V.; Salvesen, G.
J. Biol. Chem. 264, 15975-15981, 1989
A;Title: Analysis of inter-alpha-trypsin inhibitor and a novel trypsin inhibitor, pre
A;Reference number: A92736; MUID:89380192; PMID:2476436
A;Accession: B34245
A;Molecule type: protein
A;Residues: 55-74 <ENG>
R;Malki, N.; Balduyck, M.; Maes, P.; Capon, C.; Mizon, C.; Han, K.K.; Tartar, A.; Fou
Biol. Chem. Hoppe-Seyler 373, 1009-1018, 1992
A;Title: The heavy chains of human plasma inter-alpha-trypsin inhibitor: their isolat
A;Reference number: S28928; MUID:93039735; PMID:1384548
A;Accession: S28929
A;Molecule type: protein
A;Residues: 55-64 <MAL>
R;Wisniewski, H.G.; Burgess, W.H.; Oppenheim, J.D.; Vilcek, J.
Biochemistry 33, 7423-7429, 1994
A;Title: TSG-6, an arthritis-associated hyaluronan binding protein, forms a stable co
A;Reference number: A53642; MUID:94271799; PMID:7516184
A;Accession: C53642
A;Molecule type: protein
A;Residues: 55-64 <WIS>
C;Comment: Inter-alpha-trypsin inhibitor is a complex of three proteins, each derivin
C;Comment: This protein is a heterodimer of heavy and light chains.
C;Genetics:
A;Gene: GDB:ITIH2
A;Cross-references: GDB:120108; OMIM:146640
A;Map position: 10p15-10p15
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
C;Keywords: carboxyglutamic acid; glycoprotein; heterodimer; serine proteinase inhibi
F;1-18/Domain: signal sequence #status predicted <SIG>
F;19-54/Domain: propeptide #status predicted <PRO>
F;55-698/Product: inter-alpha-trypsin inhibitor heavy chain 2 #status predicted <MAT>
F;96,445/Binding site: carbohydrate (Asn) (covalent) #status experimental
F;118,671/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;282,283/Modified site: gamma-carboxyglutamic acid (Glu) #status predicted
F;421,422,423/Binding site: calcium (Asp, Gly, Asp) #status predicted

```
  Query Match             36.6%;  Score 53;  DB 1;  Length 946;
  Best Local Similarity   33.3%;  Pred. No. 15;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ::|  :  |    : | ||| | |:|
Db        212 DVWVIEPQGLRFLHVPDTFEGHFDGVP 238
```

RESULT 6
S38185
2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) ARO4 - yeast (Saccharomyces ce
N;Alternate names: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DAHP synthase;
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C;Accession: S38185; S46126; S46130; JN0322; B48651
R;Doignon, F.; Biteau, N.; Aigle, M.; Crouzet, M.
Yeast 9, 1131-1137, 1993
A;Title: The complete sequence of a 6794 bp segment located on the right arm of chromosc
A;Reference number: S38185; MUID:94078675; PMID:8256522
A;Accession: S38185
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <DOI>
A;Cross-references: GB:L20296; NID:g311101; PIDN:AAA65607.1; PID:g311102
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46126
A;Molecule type: DNA
A;Residues: 1-370 <ALJ>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR249c
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46130
A;Molecule type: DNA
A;Residues: 1-370 <AIG>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR249c
R;Kuenzler, M.; Paravicini, G.; Egli, C.M.; Irniger, S.; Braus, G.H.
Gene 113, 67-74, 1992
A;Title: Cloning, primary structure and regulation of the ARO4 gene, encoding the tyrosi
A;Reference number: JN0322; MUID:92225349; PMID:1348717
A;Accession: JN0322
A;Molecule type: DNA
A;Residues: 1-204,208-370 <KUE>
A;Cross-references: EMBL:X61107
R;Kuenzler, M.; Balmelli, T.; Egli, C.M.; Paravicini, G.; Braus, G.H.
J. Bacteriol. 175, 5548-5558, 1993
A;Title: Cloning, primary structure, and regulation of the HIS7 gene encoding a bifuncti
A;Reference number: A48651; MUID:93374850; PMID:8366040
A;Accession: B48651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 352-370 <KU2>
A;Cross-references: GB:X61107
C;Comment: This enzyme catalyzes the condensation of phosphoenolpyruvate and D-erythrose
C;Genetics:
A;Gene: SGD:ARO4
A;Cross-references: SGD:S0000453; MIPS:YBR249c
A;Map position: 2R
C;Function:
A;Description: aldehyde-lyase; carbon-carbon lyase
A;Pathway: aromatic amino acid biosynthesis; shikimate pathway
A;Note: first step in shikimate pathway
C;Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase
C;Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; cytos

```
  Query Match             35.9%;  Score 52;  DB 2;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 8.2;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEG 21
                :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100
```

RESULT 7
A42095
floral homeotic protein APETALA3 (AP3) - Arabidopsis thaliana
N;Alternate names: homeotic protein APETALA3; MADS-box regulatory protein AP3
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A42095; S52633; T47593
R;Jack, T.; Brockman, L.L.; Meyerowitz, E.M.
Cell 68, 683-697, 1992
A;Title: The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is
A;Reference number: A42095; MUID:92154682; PMID:1346756
A;Accession: A42095
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-232 <JAC>
A;Cross-references: GB:M86357; NID:g166607; PIDN:AAA32740.1; PID:g166608
A;Experimental source: petals, stamens
A;Note: sequence extracted from NCBI backbone (NCBIN:82520, NCBIP:82521)
R;Okamoto, H.; Yano, A.; Shiraishi, H.; Okada, K.; Shimura, Y.
Plant Mol. Biol. 26, 465-472, 1994
A;Title: Genetic complementation of a floral homeotic mutation, apetala3, with an Ara
A;Reference number: S52633; MUID:95036018; PMID:7948893
A;Accession: S52633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <OKA>
A;Cross-references: GB:D21125
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mew
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24469
A;Accession: T47593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <BLO>
A;Cross-references: EMBL:AL132971
A;Experimental source: cultivar Columbia; BAC clone T12E18
C;Genetics:
A;Map position: 3
A;Introns: 63/2; 85/3; 106/2; 139/3; 153/3; 168/3
A;Note: T12E18.30
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain hc
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

```
  Query Match             35.2%;  Score 51;  DB 2;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 7;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133
```

RESULT 8
C84338
spermidine/putrescine ABC transporter [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84338
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Ja
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.;
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84338
A;Status: preliminary
A;Molecule type: DNA

A;Residues: 1-374 <STO>
A;Cross-references: GB:AE004437; NID:g10581314; PIDN:AAG20071.1; GSPDB:GN00138
C;Genetics:
A;Gene: potA2

Query Match 34.5%; Score 50; DB 2; Length 374;
Best Local Similarity 76.9%; Pred. No. 16;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```
Qy         11 ELRRMSDEFEGSF 23
              ||||:||  ||||
Db        197 ELRRLSDAVEGSF 209
```

RESULT 9
A96753
probable threonine synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96753
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE005173; NID:g5903070; PIDN:AAD55628.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3N23.1
A;Map position: 1

Query Match 34.5%; Score 50; DB 2; Length 516;
Best Local Similarity 35.3%; Pred. No. 23;
Matches 12; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

```
Qy          1 NLWAAQRYGRELRRMSD--------EFEGSFKGL 26
              ||: |:|:|::  :|:|          |||| |
Db        163 NLFWAERPGKQYLQMNDLWVKHCGISHTGSFKDL 196
```

RESULT 10
A64807
endonuclease VIII (EC 3.2.-.-) - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: A64807
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A64807
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-263 <BLAT>
A;Cross-references: GB:AE000174; GB:U00096; NID:g1786920; PIDN:AAC73808.1; PID:g1786932;
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: nei
A;Map position: 16 min
C;Function:
A;Description: DNA N-glycosylase with an apurinic lyase activity; recognizes and removes
C;Superfamily: formamidopyrimidine-DNA glycosidase
C;Keywords: DNA repair; glycosidase; hydrolase; lyase

Query Match 33.8%; Score 49; DB 2; Length 263;
Best Local Similarity 47.4%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```
Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          4 GPEIRRAADNLEAAIKGKP 22
```

RESULT 11
A85572
hypothetical protein nei [imported] - Escherichia coli (strain O157:H7, substrain EDL
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 17-May-2002
C;Accession: A85572
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85572
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <STO>
A;Cross-references: GB:AE005174; NID:g12513625; PIDN:AAG55037.1; GSPDB:GN00145; UWGP:
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: nei
C;Superfamily: formamidopyrimidine-DNA glycosidase

Query Match 33.8%; Score 49; DB 2; Length 263;
Best Local Similarity 47.4%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```
Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          4 GPEIRRAADNLEAAIKGKP 22
```

RESULT 12
C90721
hypothetical protein ECs0739 [imported] - Escherichia coli (strain O157:H7, substrain
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 17-May-2002
C;Accession: C90721
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C90721
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB34162.1; PID:g13360197; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0739
C;Superfamily: formamidopyrimidine-DNA glycosidase

Query Match 33.8%; Score 49; DB 2; Length 263;
Best Local Similarity 47.4%; Pred. No. 16;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

```
Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          4 GPEIRRAADNLEAAIKGKP 22
```

RESULT 13
AD0590
endonuclease VIII, DNA N-glycosylase with an AP lyase activity STY0771 [imported] - S

C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AD0590
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AD0590
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-263 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD05190.1; PID:g16501960; GSPDB:GN00176
C;Genetics:
A;Gene: STY0771
C;Superfamily: formamidopyrimidine-DNA glycosidase

  Query Match 33.8%; Score 49; DB 2; Length 263;
  Best Local Similarity 47.4%; Pred. No. 16;
  Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

Qy 9 GRELRRMSDEFEGSFKGLP 27
Db 4 GPEIRRAADNLEAAIKGKP 22

RESULT 14
E83517
conserved hypothetical protein PA1031 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83517
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathc
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <STO>
A;Cross-references: GB:AE004535; GB:AE004091; NID:g9946936; PIDN:AAG04420.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1031

  Query Match 33.8%; Score 49; DB 2; Length 453;
  Best Local Similarity 55.6%; Pred. No. 28;
  Matches 10; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

Qy 3 WAAQRYGR--ELRRMSDE 18
Db 65 WASERQGREEELRRLASE 82

RESULT 15
S40376
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40376
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40376
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-134 <KLE>
A;Cross-references: EMBL:X72486; NID:g441440; PIDN:CAA51154.1; PID:g441441
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-113/Domain: immunoglobulin homology <IMM>

  Query Match 33.4%; Score 48.5; DB 2; Length 134;
  Best Local Similarity 38.2%; Pred. No. 9.3;
  Matches 13; Conservative 1; Mismatches 9; Indels 11; Gaps 1;

Qy 3 WAAQRYGRELRRM----------SDEFEGSFKG 25
Db 58 WFRQRPGRSPRRLIYNVSKRDSGVSDRFSGSGSG 91

Search completed: September 15, 2003, 17:27:01
Job time : 13.15 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 6.36429 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	145	100.0	204	1	BAD_MOUSE	Q61337 mus musculu
2	145	100.0	205	1	BAD_RAT	O35147 rattus norv
3	120.5	83.1	168	1	BAD_HUMAN	Q92934 homo sapien
4	61	42.1	946	1	ITH2_MESAU	P97279 mesocricetu
5	60	41.4	946	1	ITH2_MOUSE	Q61703 mus musculu
6	53	36.6	946	1	ITH2_HUMAN	P19823 homo sapien
7	52	35.9	370	1	AROG_YEAST	P32449 saccharomyc
8	51	35.2	232	1	AP3_ARATH	P35632 arabidopsis
9	50	34.5	851	1	CE05_MOUSE	Q8k2h3 mus musculu
10	49.5	34.1	506	1	MATK_LEDPA	O62992 ledum palus
11	49.5	34.1	506	1	MATK_RHOFR	O62984 rhododendro
12	49.5	34.1	506	1	MATK_RHOTS	O62991 rhododendro
13	49	33.8	262	1	END8_ECO57	Q8x9c6 escherichia
14	49	33.8	262	1	END8_ECOLI	P50465 escherichia
15	49	33.8	262	1	END8_SALTI	Q8z8d2 salmonella
16	49	33.8	262	1	END8_SALTY	Q8zqu6 salmonella
17	49	33.8	453	1	RMUC_PSEAE	Q9i4u3 pseudomonas
18	48	33.1	205	1	RAS3_RHIRA	P22280 rhizomucor
19	48	33.1	220	1	6PGL_THEMA	Q9x0n8 thermotoga
20	48	33.1	519	1	THRC_SOLTU	Q9mt28 solanum tub
21	48	33.1	526	1	THRC_ARATH	Q9s7b5 arabidopsis
22	47	32.4	198	1	BIM_HUMAN	O43521 homo sapien
23	46.5	32.1	429	1	FMR2_ANTEL	Q16994 anthopleura
24	46.5	32.1	435	1	FMR1_ANTEL	P10419 anthopleura
25	46.5	32.1	855	1	XAB2_MOUSE	Q9dcd2 mus musculu
26	46.5	32.1	855	1	XAB2_RAT	Q99pk0 rattus norv
27	46	31.7	738	1	SEC6_DROME	Q9v8k2 drosophila
28	46	31.7	1378	1	RPOB_CAMJE	Q46124 campylobact
29	45.5	31.4	287	1	PRFA_POLPE	P21259 polyorchis
30	45.5	31.4	334	1	FMRA_CALPA	Q01133 calliactis
31	45.5	31.4	507	1	MATK_LOIPR	O47169 loiseleuria
32	45.5	31.4	855	1	XAB2_HUMAN	Q9hcs7 homo sapien
33	45	31.0	328	1	SNF4_KLULA	Q9p869 kluyveromyc
34	45	31.0	513	1	KCG1_HUMAN	Q9uix4 homo sapien
35	45	31.0	590	1	DCOA_SALTY	Q03030 salmonella
36	45	31.0	595	1	DCOA_KLEPN	P13187 klebsiella
37	45	31.0	653	1	HT2A_HUMAN	Q13049 homo sapien
38	45	31.0	865	1	ENV_SIVAT	P05886 simian immu
39	45	31.0	1535	1	LML1_CAEEL	Q18823 caenorhabdi
40	44.5	30.7	143	1	NB6M_HUMAN	Q9p0j0 h nadh-ubiq
41	44.5	30.7	907	1	NUOG_ECOLI	P33602 escherichia
42	44.5	30.7	907	1	NUOG_SALTY	P33900 salmonella
43	44	30.3	196	1	BIM_MOUSE	O54918 mus musculu
44	44	30.3	196	1	BIM_RAT	O88498 rattus norv
45	44	30.3	768	1	ENV_SIVA1	P27757 simian immu

ALIGNMENTS

RESULT 1
BAD_MOUSE
ID BAD_MOUSE STANDARD; PRT; 204 AA.
AC Q61337;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE 6) (Bcl-xL/Bcl-2 associated death promoter).
GN BAD OR BBC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Thymus;
RX MEDLINE=95136361; PubMed=7834748;
RA Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT promotes cell death.";
RL Cell 80:285-291(1995).
RN [2]
RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX MEDLINE=98022383; PubMed=9381178;
RA Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT "Interleukin-3-induced phosphorylation of BAD through the protein
RT kinase Akt.";
RL Science 278:687-689(1997).
RN [3]
RP MUTAGENESIS OF SERINE RESIDUES.
RX MEDLINE=20403302; PubMed=10949026;
RA Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA Greenberg M.E.;
RT "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT BH3 domain phosphorylation.";
RL Mol. Cell 6:41-51(2000).
CC -!- FUNCTION: Promotes cell death. Successfully competes for the
CC binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC of heterodimerization of these proteins with BAX. Can reverse the
CC death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC Appears to act as a link between growth factor receptor signaling
CC and the apoptotic pathways.
CC -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC phosphorylation, locates to the cytoplasm.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC with 14-3-3 proteins. This interaction then facilitates the
CC phosphorylation at Ser-155, a site within the BH3 domain, leading
CC to the release of Bcl-x(L) and the promotion of cell survival.

```
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             100.0%;  Score 145;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGLP 166


RESULT 2
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;
```

```
  Query Match             100.0%;  Score 145;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 1.4e-14;
  Matches   27;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |||||||||||||||||||||||||||
Db        141 NLWAAQRYGRELRRMSDEFEGSFKGLP 167


RESULT 3
BAD_HUMAN
ID   BAD_HUMAN      STANDARD;      PRT;   168 AA.
AC   Q92934; O14803;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE   XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN   BAD OR BBC6 OR BCL2L8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT   "A human protein that interacts with Bcl-2 and have homology to mouse
RT   BAD.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX   MEDLINE=97083574; PubMed=8929532;
RA   Wang H.-G., Rapp U.R., Reed J.C.;
RT   "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL   Cell 87:629-638(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Takayama S., Reed J.C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND DIMERIZATION.
RC   TISSUE=Bone marrow;
RX   MEDLINE=98049554; PubMed=9388232;
RA   Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA   Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT   "Dimerization properties of human BAD.";
RL   J. Biol. Chem. 272:30866-30872(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match             83.1%;  Score 120.5;  DB 1;  Length 168;
```

```
  Best Local Similarity   89.3%;  Pred. No. 5.5e-11;
  Matches   25;  Conservative    0;  Mismatches    2;  Indels    1;  Gaps    1;

Qy          1 NLWAAQRYGRELRRMSDEFEGSF-KGLP 27
              |||||||||||||||||||  || ||||
Db        103 NLWAAQRYGRELRRMSDEFVDSFKKGLP 130


RESULT 4
ITH2_MESAU
ID   ITH2_MESAU     STANDARD;      PRT;   946 AA.
AC   P97279;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2) (HC2).
GN   ITIH2.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97420688; PubMed=9276673;
RA   Nakatani T., Suzuki Y., Yamamoto T., Sinohara H.;
RT   "Molecular cloning and sequencing of cDNAs encoding three heavy-chain
RT   precursors of the inter-alpha-trypsin inhibitor in Syrian hamster:
RT   implications for the evolution of the inter-alpha-trypsin inhibitor
RT   heavy chain family.";
RL   J. Biochem. 122:71-82(1997).
RN   [2]
RP   SEQUENCE OF 55-64; 140-146; 151-156; 424-447; 500-528 AND 577-605,
RP   AND SUBUNITS.
RC   TISSUE=Plasma;
RX   MEDLINE=97018241; PubMed=8864857;
RA   Yamamoto T., Yamamoto K., Sinohara H.;
RT   "Inter-alpha-trypsin inhibitor and its related proteins in Syrian
RT   hamster urine and plasma.";
RL   J. Biochem. 120:145-152(1996).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89286; BAA13939.1; -.
DR   PIR; JC5575; JC5575.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL       1     18       POTENTIAL.
FT   PROPEP      19     54       BY SIMILARITY.
FT   CHAIN       55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                               H2.
FT   PROPEP     703    946       BY SIMILARITY.
FT   DOMAIN     308    468       VWFA.
FT   CARBOHYD   118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING    702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                               (BY SIMILARITY).
FT   CONFLICT   510    510       V -> Y (IN REF. 2).
FT   CONFLICT   595    595       E -> I (IN REF. 2).
SQ   SEQUENCE   946 AA;  106580 MW;  CA8BF565458E7B2E CRC64;

  Query Match             42.1%;  Score 61;  DB 1;  Length 946;
  Best Local Similarity   37.0%;  Pred. No. 0.34;
  Matches   10;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |:|  :  |     : | ||| |:|:|
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGVP 238


RESULT 5
ITH2_MOUSE
ID   ITH2_MOUSE     STANDARD;      PRT;   946 AA.
AC   Q61703;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2).
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6N; TISSUE=Liver;
RX   MEDLINE=95194326; PubMed=7534067;
RA   Chan P., Risler J.-L., Raguenez G., Salier J.-P.;
RT   "The three heavy-chain precursors for the inter-alpha-inhibitor
RT   family in mouse: new members of the multicopper oxidase protein group
RT   with differential transcription in liver and brain.";
RL   Biochem. J. 306:505-512(1995).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH LIVER AND BRAIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X70392; CAA49842.1; -.
DR   PIR; S54354; S54354.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                (BY SIMILARITY).
SQ   SEQUENCE   946 AA;  105927 MW;  40DB6716433ED9DC CRC64;

  Query Match             41.4%;  Score 60;  DB 1;  Length 946;
  Best Local Similarity   37.0%;  Pred. No. 0.47;
  Matches   10;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |:|   :  |     : | ||| |:|:|
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGVP 238


RESULT 6
ITH2_HUMAN
ID   ITH2_HUMAN     STANDARD;      PRT;   946 AA.
AC   P19823; Q14659; Q15484;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2) (Inter-alpha-trypsin
DE   inhibitor complex component II) (Serum-derived hyaluronan-associated
DE   protein) (SHAP).
GN   ITIH2 OR IGHEP2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX   MEDLINE=88152237; PubMed=2450046;
RA   Gebhard W., Schreitmueller T., Hochstrasser K., Wachter E.;
RT   "Complementary DNA and derived amino acid sequence of the precursor
RT   of one of the three protein components of the inter-alpha-trypsin
RT   inhibitor complex.";
RL   FEBS Lett. 229:63-67(1988).
RN   [2]
RP   SEQUENCE OF 384-865 FROM N.A.
RX   MEDLINE=88068576; PubMed=2446322;
RA   Salier J.-P., Diarra-Mehrpour M., Sesboue R., Bourguignon J.,
RA   Benarous R., Ohkubo I., Kurachi S., Kurachi K., Martin J.P.;
RT   "Isolation and characterization of cDNAs encoding the heavy chain of
RT   human inter-alpha-trypsin inhibitor (I alpha TI): unambiguous
RT   evidence for multipolypeptide chain structure of I alpha TI.";
RL   Proc. Natl. Acad. Sci. U.S.A. 84:8272-8276(1987).
RN   [3]
RP   SEQUENCE OF 384-766 FROM N.A.
RX   MEDLINE=89076497; PubMed=2462430;
RA   Salier J.-P., Diarra-Mehrpour M., Sesboue R., Bourguignon J.,
RA   Martin J.P.;
RT   "Human inter-alpha-trypsin inhibitor. Isolation and characterization
RT   of heavy (H) chain cDNA clones coding for a 383 amino-acid sequence
RT   of the H chain.";
RL   Biol. Chem. Hoppe-Seyler 369:15-18(1988).
RN   [4]
RP   PARTIAL SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=88024442; PubMed=3663330;
RA   Schreitmueller T., Hochstrasser K., Resinger P.W.M., Wachter E.,
RA   Gebhard W.;
RT   "cDNA cloning of human inter-alpha-trypsin inhibitor discloses three
RT   different proteins.";
RL   Biol. Chem. Hoppe-Seyler 368:963-970(1987).
RN   [5]
RP   SEQUENCE OF 55-74; 116-127; 224-246; 295-307 AND 365-385.
RX   MEDLINE=89380192; PubMed=2476436;
RA   Enghild J.J., Thoegersen I.B., Pizzo S.V., Salvesen G.;
RT   "Analysis of inter-alpha-trypsin inhibitor and a novel trypsin
RT   inhibitor, pre-alpha-trypsin inhibitor, from human plasma.
RT   Polypeptide chain stoichiometry and assembly by glycan.";
RL   J. Biol. Chem. 264:15975-15981(1989).
RN   [6]
RP   SEQUENCE OF 55-64.
RC   TISSUE=Plasma;
RX   MEDLINE=93039735; PubMed=1384548;
RA   Malki N., Balduyck M., Maes P., Capon C., Mizon C., Han K.K.,
RA   Tartar A., Fournet B., Mizon J.;
RT   "The heavy chains of human plasma inter-alpha-trypsin inhibitor: their
RT   isolation, their identification by electrophoresis and partial
RT   sequencing. Differential reactivity with concanavalin A.";
RL   Biol. Chem. Hoppe-Seyler 373:1009-1018(1992).
RN   [7]
RP   SEQUENCE OF 55-64 AND 681-702, CROSS-LINK STRUCTURE, AND
RP   CARBOHYDRATE-LINKAGE SITES THR-691.
RX   MEDLINE=93232026; PubMed=7682553;
RA   Enghild J.J., Salvesen G., Thoegersen I.B., Valnickova Z.,
RA   Pizzo S.V., Hefta S.A.;
RT   "Presence of the protein-glycosaminoglycan-protein covalent cross-link
RT   in the inter-alpha-inhibitor-related proteinase inhibitor heavy chain
RT   2/bikunin.";
RL   J. Biol. Chem. 268:8711-8716(1993).
RN   [8]
RP   SEQUENCE OF 67-101, AND HYALURONAN BINDING.
RC   TISSUE=Serum;
RX   MEDLINE=94075371; PubMed=7504674;
RA   Huang L., Yoneda M., Kimata K.;
RT   "A serum-derived hyaluronan-associated protein (SHAP) is the heavy
RT   chain of the inter alpha-trypsin inhibitor.";
RL   J. Biol. Chem. 268:26725-26730(1993).
RN   [9]
RP   SEQUENCE OF 699-702, AND COVALENT LINKAGE WITH CHONDROITIN SULFATE.
RC   TISSUE=Plasma;
RX   MEDLINE=94229087; PubMed=7513643;
RA   Morelle W., Capon C., Balduyck M., Sautiere P., Kouach M.,
RA   Michalski C., Fournet B., Mizon J.;
RT   "Chondroitin sulphate covalently cross-links the three polypeptide
RT   chains of inter-alpha-trypsin inhibitor.";
RL   Eur. J. Biochem. 221:881-888(1994).
RN   [10]
RP   CARBOHYDRATE-LINKAGE SITES, AND MASS SPECTROMETRY.
RX   MEDLINE=98343966; PubMed=9677337;
RA   Flahaut C., Capon C., Balduyck M., Ricart G., Sautiere P., Mizon J.;
RT   "Glycosylation pattern of human inter-alpha-inhibitor heavy chains.";
RL   Biochem. J. 333:749-756(1998).
RN   [11]
RP   CARBOHYDRATE-LINKAGE SITES, AND DISULFIDE BONDS.
RX   MEDLINE=98087700; PubMed=9425062;
RA   Olsen E.H.N., Rahbek-Nielsen H., Thoegersen I.B., Roepstorff P.,
RA   Enghild J.J.;
RT   "Posttranslational modifications of human inter-alpha-inhibitor:
```

```
RT   identification of glycans and disulfide bridges in heavy chains 1 and
RT   2.";
RL   Biochemistry 37:408-416(1998).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES.
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE.
CC   -!- MASS SPECTROMETRY: MW=76508; METHOD=MALDI; RANGE=55-702.
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X07173; CAA30160.1; ALT_SEQ.
DR   EMBL; M18193; AAA60558.1; -.
DR   EMBL; M33033; AAA59195.1; -.
DR   PIR; S00346; IYHU2.
DR   GlycoSuiteDB; P19823; -.
DR   Genew; HGNC:6167; ITIH2.
DR   MIM; 146640; -.
DR   GO; GO:0005211; F:plasma glycoprotein; TAS.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Gamma-carboxyglutamic acid; Glycoprotein.
FT   SIGNAL       1    18       POTENTIAL.
FT   PROPEP      19    54
FT   CHAIN       55   702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                               H2.
FT   PROPEP     703   946
FT   DOMAIN     308   468       VWFA.
FT   DISULFID   261   264
FT   DISULFID   650   651
FT   CARBOHYD   118   118       N-LINKED (GLCNAC. . .) (COMPLEX).
FT                               /FTId=CAR_000140.
FT   CARBOHYD   666   666       O-LINKED (GALNAC. . .) (PARTIAL).
FT                               /FTId=CAR_000214.
FT   CARBOHYD   671   671       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   673   673       O-LINKED (GALNAC. . .).
FT                               /FTId=CAR_000215.
FT   CARBOHYD   675   675       O-LINKED (GALNAC. . .).
FT                               /FTId=CAR_000216.
FT   CARBOHYD   691   691       O-LINKED (GALNAC. . .).
FT                               /FTId=CAR_000217.
FT   MOD_RES    282   282       GAMMA-CARBOXYGLUTAMIC ACID.
FT   MOD_RES    283   283       GAMMA-CARBOXYGLUTAMIC ACID.
FT   BINDING    702   702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE.
FT   CONFLICT   374   374       K -> L (IN REF. 5).
FT   CONFLICT   674   674       P -> A (IN REF. 2 AND 3).
FT   CONFLICT   705   705       F -> S (IN REF. 2 AND 3).
FT   CONFLICT   729   729       N -> D (IN REF. 2 AND 3).
FT   CONFLICT   731   731       V -> A (IN REF. 2 AND 3).
SQ   SEQUENCE   946 AA;  106436 MW;  1478CF3E8F3BA776 CRC64;

  Query Match           36.6%;  Score 53;  DB 1;  Length 946;
  Best Local Similarity   33.3%;  Pred. No. 5.4;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ::|  :  |     : | ||| | |:|
Db        212 DVWVIEPQGLRFLHVPDTFEGHFDGVP 238


RESULT 7
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
```

```
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match             35.9%;  Score 52;  DB 1;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 2.7;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEG 21
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 8
AP3_ARATH
ID   AP3_ARATH      STANDARD;      PRT;   232 AA.
AC   P35632; Q39003; Q8LB79; Q9S7Q3; Q9SQ14; Q9SQ15; Q9SQ16; Q9SQ17;
AC   Q9SQ18; Q9SQ19; Q9SQ20; Q9SQ21; Q9SQ22; Q9SX13;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Floral homeotic protein APETALA3.
GN   AP3 OR AT3G54340 OR T12E18_30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Petal;
RX   MEDLINE=92154682; PubMed=1346756;
RA   Jack T., Brockman L.L., Meyerowitz E.M.;
RT   "The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS
RT   box and is expressed in petals and stamens.";
RL   Cell 68:683-697(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Landsberg erecta;
RX   MEDLINE=95036018; PubMed=7948893;
RA   Okamoto H., Yano A., Shiraishi H., Okada K., Shimura Y.;
RT   "Genetic complementation of a floral homeotic mutation, apetala3,
RT   with an Arabidopsis thaliana gene homologous to DEFICIENS of
RT   Antirrhinum majus.";
RL   Plant Mol. Biol. 26:465-472(1994).
RN   [3]
RP   SEQUENCE FROM N.A., AND VARIANTS.
RC   STRAIN=cv. Bla-1, cv. Bretagny, cv. Bs-1, cv. Bu-0, cv. Bu-2,
RC   cv. Chi-1, cv. Co-1, cv. Columbia, cv. Corsacalla-1, cv. Cvi-0,
RC   cv. Gr-3, cv. Jl-1, cv. Kas-1, cv. Kent, cv. Landsberg erecta,
RC   cv. Li-3, cv. Li-8, and cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci: departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=21016720; PubMed=11130713;
RA   Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA   Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA   Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA   De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA   Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA   Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA   Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA   Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA   Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA   Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA   Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA   Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA   Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA   de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA   Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA   Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA   Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA   Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA   Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA   Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA   Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA   Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA   Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA   Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA   Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT   "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT   thaliana.";
RL   Nature 408:820-822(2000).
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT   "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT   SSP consortium (Salk/Stanford/PGEC).";
RL   Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE OF 36-128 FROM N.A.
RC   STRAIN=cv. Columbia;
RX   MEDLINE=99311297; PubMed=10382288;
RA   Brunel D., Froger N., Pelletier G.;
RT   "Development of amplified consensus genetic markers (ACGM) in Brassica
RT   napus from Arabidopsis thaliana sequences of known biological
RT   function.";
RL   Genome 42:387-402(1999).
RN   [8]
RP   FUNCTION.
RX   PubMed=8565821;
RA   Krizek B.A., Meyerowitz E.M.;
RT   "The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient
RT   to provide the B class organ identity function.";
RL   Development 122:11-22(1996).
RN   [9]
RP   CHARACTERIZATION.
RX   PubMed=8643482;
RA   Riechmann J.L., Krizek B.A., Meyerowitz E.M.;
RT   "Dimerization specificity of Arabidopsis MADS domain homeotic proteins
RT   APETALA1, APETALA3, PISTILLATA, and AGAMOUS.";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:4793-4798(1996).
RN   [10]
RP   GENETIC REGULATION.
RX   PubMed=11283333;
RA   Ng M., Yanofsky M.F.;
RT   "Activation of the Arabidopsis B class homeotic genes by APETALA1.";
RL   Plant Cell 13:739-753(2001).
RN   [11]
RP   CHARACTERIZATION.
RX   PubMed=11206550;
RA   Honma T., Goto K.;
RT   "Complexes of MADS-box proteins are sufficient to convert leaves into
RT   floral organs.";
RL   Nature 409:525-529(2001).
```

```
CC   -!- FUNCTION: Probable transcription factor involved in the genetic
CC       control of flower development. Is required for normal development
CC       of petals and stamens in the wild-type flower. Forms an
CC       heterodimer with PISTILLATA that is required for autoregulation of
CC       both AP3 and PI genes. AP3/PI heterodimer interacts with APETALA1
CC       or SEPALLATA3 to form a ternary complex that could be responsible
CC       for the regulation of the genes involved in the flower
CC       development.
CC   -!- SUBUNIT: Forms an heterodimer with PISTILLATA, capable of binding
CC       to CArG-box sequences. AP3/PI heterodimer binds AP1 or SEP3 to
CC       form complexes.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Expressed in petals and stamens.
CC   -!- INDUCTION: Positively regulated by the meristem identity proteins
CC       APETALA1 and LEAFY with the cooperation of UFO.
CC   -!- MISCELLANEOUS: Mutations in AP3 cause transformation of petals
CC       into sepals and stamina into carpels.
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
CC   -!- SIMILARITY: Contains 1 K-box dimerization domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M86357; AAA32740.1; -.
DR   EMBL; D21125; BAA04665.1; -.
DR   EMBL; AF115798; AAD51887.1; -.
DR   EMBL; AF115799; AAD51888.1; -.
DR   EMBL; AF115800; AAD51889.1; -.
DR   EMBL; AF115801; AAD51890.1; -.
DR   EMBL; AF115802; AAD51891.1; -.
DR   EMBL; AF115803; AAD51892.1; -.
DR   EMBL; AF115804; AAD51893.1; -.
DR   EMBL; AF115805; AAD51894.1; -.
DR   EMBL; AF115806; AAD51895.1; -.
DR   EMBL; AF115807; AAD51896.1; -.
DR   EMBL; AF115808; AAD51897.1; -.
DR   EMBL; AF115809; AAD51898.1; -.
DR   EMBL; AF115810; AAD51899.1; -.
DR   EMBL; AF115811; AAD51900.1; -.
DR   EMBL; AF115812; AAD51901.1; -.
DR   EMBL; AF115813; AAD51902.1; -.
DR   EMBL; AF115814; AAD51903.1; -.
DR   EMBL; AL132971; CAB81799.1; -.
DR   EMBL; AY087369; AAM64919.1; -.
DR   EMBL; AY070397; AAL49893.1; -.
DR   EMBL; AY142590; AAN13159.1; -.
DR   EMBL; AF056541; AAD41557.1; -.
DR   PIR; A42095; A42095.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T01776; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Flowering; Transcription regulation; Activator; Developmental protein;
KW   Nuclear protein; DNA-binding; Coiled coil; Polymorphism.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       93    165       K-BOX.
FT   DOMAIN       75    164       COILED COIL (POTENTIAL).
FT   VARIANT      31     31       K -> R (in strain cv. Lisse).
FT   VARIANT      47     47       M -> T (in strain cv. Bretagny).
FT   VARIANT      61     61       N -> D (in strain cv. Corsacalla-1).
FT   VARIANT      73     73       T -> S (in strain cv. Li-8).
FT   VARIANT     109    109       L -> V (in strain cv. Kas-1).
FT   VARIANT     115    115       E -> K (in strains cv. Chi-1 and cv.
FT                                Gr-3).

  Query Match             35.2%;  Score 51;  DB 1;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 2.3;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 9
CE05_MOUSE
ID   CE05_MOUSE     STANDARD;      PRT;   851 AA.
AC   Q8K2H3;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Protein C5orf5 homolog.
GN   C5ORF5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Belongs to the FAM13 family.
CC   -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC031465; AAH31465.1; -.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   GTPase activation.
FT   DOMAIN       23    212       RHO-GAP.
FT   DOMAIN      189    256       GLU-RICH.
SQ   SEQUENCE   851 AA;  97054 MW;  C2B26669FB6DB2CE CRC64;

  Query Match             34.5%;  Score 50;  DB 1;  Length 851;
  Best Local Similarity   45.5%;  Pred. No. 13;
```

```
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          2 LWAAQRYGRELRRMSDEFEGSF 23
              || |:   ::||:|  ||| :|
Db        784 LWKARAEKKKLRKMLREFEEAF 805


RESULT 10
MATK_LEDPA
ID   MATK_LEDPA     STANDARD;      PRT;   506 AA.
AC   O62992;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Ledum palustre (Wild rosemary).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75583;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012751; BAA25872.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60412 MW;  CFEA926307DAC85E CRC64;

  Query Match             34.1%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 9;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 11
MATK_RHOFR
ID   MATK_RHOFR     STANDARD;      PRT;   506 AA.
AC   O62984;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron ferrugineum (Alpenrose).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49622;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012741; BAA25862.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60534 MW;  ADA44B25E92436E8 CRC64;

  Query Match             34.1%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 9;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 12
MATK_RHOTS
ID   MATK_RHOTS     STANDARD;      PRT;   506 AA.
AC   O62991;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron tsusiophyllum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49629;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
```

```
DR   EMBL; AB012750; BAA25871.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60569 MW;  AEE12FF8809C223E CRC64;

  Query Match             34.1%;  Score 49.5;  DB 1;  Length 506;
  Best Local Similarity   36.7%;  Pred. No. 9;
  Matches   11;  Conservative    5;  Mismatches    7;  Indels    7;  Gaps    1;

Qy          2 LWAA-------QRYGRELRRMSDEFEGSFK 24
              :|||       :|:||  | :|  : || |
Db        393 VWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 13
END8_ECO57
ID   END8_ECO57     STANDARD;      PRT;   262 AA.
AC   Q8X9C6;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Endonuclease VIII (EC 3.2.-.-).
GN   NEI OR Z0865 OR ECS0739.
OS   Escherichia coli O157:H7.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=83334;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX   MEDLINE=21074935; PubMed=11206551;
RA   Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA   Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA   Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA   Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA   Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA   Welch R.A., Blattner F.R.;
RT   "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL   Nature 409:529-533(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=O157:H7 / RIMD 0509952;
RX   MEDLINE=21156231; PubMed=11258796;
RA   Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA   Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
RA   Iida T., Takami H., Honda T., Sasakawa C., Ogasawara N., Yasunaga T.,
RA   Kuhara S., Shiba T., Hattori M., Shinagawa H.;
RT   "Complete genome sequence of enterohemorrhagic Escherichia coli
RT   O157:H7 and genomic comparison with a laboratory strain K-12.";
RL   DNA Res. 8:11-22(2001).
CC   -!- FUNCTION: DNA N-GLYCOSYLASE WITH AN AP LYASE ACTIVITY. REQUIRED
CC       FOR THE REPAIR OF OXIDATIVE DNA DAMAGE. CLEAVES THE DNA BACKBONE
CC       BY BETA-DELTA ELIMINATION AS WELL AS 5'DEOXYRIBOSE PHOSPHATE (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE FPG FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE005249; AAG55037.1; -.
DR   EMBL; AP002552; BAB34162.1; -.
DR   PIR; A85572; A85572.
DR   PIR; C90721; C90721.
DR   InterPro; IPR000191; Fapy_DNA_glyco.
DR   InterPro; IPR000214; Fapy_DNAglyco_zn.
DR   Pfam; PF01149; Fapy_DNA_glyco; 1.
DR   ProDom; PD003680; Fapy_DNA_glyco; 1.
DR   PROSITE; PS01242; FPG; 1.
KW   DNA repair; Hydrolase; Glycosidase; Endonuclease; Zinc; Zinc-finger;
KW   Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   ZN_FING     237    260       POTENTIAL.
SQ   SEQUENCE   262 AA;  29680 MW;  37C3C5E236E07A88 CRC64;

  Query Match             33.8%;  Score 49;  DB 1;  Length 262;
  Best Local Similarity   47.4%;  Pred. No. 5.2;
  Matches    9;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          3 GPEIRRAADNLEAAIKGKP 21


RESULT 14
END8_ECOLI
ID   END8_ECOLI     STANDARD;      PRT;   262 AA.
AC   P50465;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Endonuclease VIII (EC 3.2.-.-).
GN   NEI OR B0714.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A., AND SEQUENCE OF 1-35; 188-205 AND 213-226.
RX   MEDLINE=97315255; PubMed=9171429;
RA   Jiang D., Hatahet Z., Blaisdell J.O., Melamede R.J., Wallace S.S.;
RT   "Escherichia coli endonuclease VIII: cloning, sequencing, and
RT   overexpression of the nei structural gene and characterization of nei
RT   and nei nth mutants.";
RL   J. Bacteriol. 179:3773-3782(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97315256; PubMed=9171430;
RA   Saito Y., Uraki F., Nakajima S., Asaeda A., Ono K., Kubo K.,
RA   Yamamoto K.;
RT   "Characterization of endonuclease III (nth) and endonuclease VIII
RT   (nei) mutants of Escherichia coli K-12.";
RL   J. Bacteriol. 179:3783-3785(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Mori H.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: DNA N-GLYCOSYLASE WITH AN AP LYASE ACTIVITY. REQUIRED
CC       FOR THE REPAIR OF OXIDATIVE DNA DAMAGE. CLEAVES THE DNA BACKBONE
CC       BY BETA-DELTA ELIMINATION AS WELL AS 5'DEOXYRIBOSE PHOSPHATE.
CC   -!- SIMILARITY: BELONGS TO THE FPG FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ----------------------------------------------------------------------
DR   EMBL; U38616; AAC45355.1; -.
DR   EMBL; D89754; BAA20414.1; -.
DR   EMBL; AE000174; AAC73808.1; -.
DR   EMBL; D90710; BAA35378.1; -.
DR   PIR; A64807; A64807.
DR   PDB; 1K3W; 04-OCT-02.
DR   EcoGene; EG13237; nei.
DR   InterPro; IPR000191; Fapy_DNA_glyco.
DR   InterPro; IPR000214; Fapy_DNAglyco_zn.
DR   Pfam; PF01149; Fapy_DNA_glyco; 1.
DR   ProDom; PD003680; Fapy_DNA_glyco; 1.
DR   PROSITE; PS01242; FPG; 1.
KW   DNA repair; Hydrolase; Glycosidase; Endonuclease; Zinc; Zinc-finger;
KW   Complete proteome; 3D-structure.
FT   INIT_MET      0      0
FT   ZN_FING     237    260       POTENTIAL.
SQ   SEQUENCE   262 AA;  29714 MW;  5010961768ADC265 CRC64;

  Query Match             33.8%;  Score 49;  DB 1;  Length 262;
  Best Local Similarity   47.4%;  Pred. No. 5.2;
  Matches    9;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          3 GPEIRRAADNLEAAIKGKP 21


RESULT 15
END8_SALTI
ID   END8_SALTI     STANDARD;      PRT;   262 AA.
AC   Q8Z8D2;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Endonuclease VIII (EC 3.2.-.-).
GN   NEI OR STY0771 OR T2148.
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
CC   -!- FUNCTION: DNA N-GLYCOSYLASE WITH AN AP LYASE ACTIVITY. REQUIRED
CC       FOR THE REPAIR OF OXIDATIVE DNA DAMAGE. CLEAVES THE DNA BACKBONE
CC       BY BETA-DELTA ELIMINATION AS WELL AS 5'DEOXYRIBOSE PHOSPHATE (By
CC       similarity).
CC   -!- SIMILARITY: BELONGS TO THE FPG FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AL627268; CAD05190.1; -.
DR   EMBL; AE016841; AAO69762.1; -.
DR   InterPro; IPR000191; Fapy_DNA_glyco.
DR   InterPro; IPR000214; Fapy_DNAglyco_zn.
DR   Pfam; PF01149; Fapy_DNA_glyco; 1.
DR   ProDom; PD003680; Fapy_DNA_glyco; 1.
DR   PROSITE; PS01242; FPG; 1.
KW   DNA repair; Hydrolase; Glycosidase; Endonuclease; Zinc; Zinc-finger;
KW   Complete proteome.
FT   INIT_MET      0      0       BY SIMILARITY.
FT   ZN_FING     237    260       POTENTIAL.
SQ   SEQUENCE   262 AA;  29734 MW;  4E255C0CCF59A6A3 CRC64;

  Query Match             33.8%;  Score 49;  DB 1;  Length 262;
  Best Local Similarity   47.4%;  Pred. No. 5.2;
  Matches    9;  Conservative    3;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          9 GRELRRMSDEFEGSFKGLP 27
              | |:|| :|  | : || |
Db          3 GPEIRRAADNLEAAIKGKP 21


Search completed: September 15, 2003, 17:22:59
Job time : 6.36429 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 29.3143 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-3
Perfect score: 145
Sequence: 1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	60.0	146	13	Q9I9N2	Q9i9n2 brachydanio
2	60	41.4	946	11	Q8K016	Q8k016 mus musculu
3	53	36.6	196	16	Q8VJS3	Q8vjs3 mycobacteri
4	53	36.6	223	16	Q10843	Q10843 mycobacteri
5	52	35.9	471	17	Q8ZY71	Q8zy71 pyrobaculum
6	52	35.9	946	6	Q9GLY6	Q9gly6 oryctolagus
7	51	35.2	231	10	Q9SEG0	Q9seg0 arabidopsis
8	51	35.2	232	10	Q9SQ20	Q9sq20 arabidopsis
9	51	35.2	232	10	Q9SQ22	Q9sq22 arabidopsis
10	51	35.2	232	10	Q9SQ17	Q9sq17 arabidopsis
11	51	35.2	232	10	Q9SQ19	Q9sq19 arabidopsis
12	51	35.2	232	10	Q9SQ21	Q9sq21 arabidopsis
13	51	35.2	232	10	Q8LB79	Q8lb79 arabidopsis
14	51	35.2	232	10	Q9SQ16	Q9sq16 arabidopsis
15	51	35.2	232	10	Q9SQ15	Q9sq15 arabidopsis
16	51	35.2	232	10	Q9S7Q3	Q9s7q3 arabidopsis
17	51	35.2	232	10	Q9SQ18	Q9sq18 arabidopsis
18	50.5	34.8	904	2	Q9KGW3	Q9kgw3 pseudomonas
19	50.5	34.8	909	16	Q8EI34	Q8ei34 shewanella
20	50.5	34.8	990	10	Q9C793	Q9c793 arabidopsis
21	50	34.5	168	11	Q8K316	Q8k316 mus musculu
22	50	34.5	283	15	O37056	O37056 chimpanzee
23	50	34.5	374	17	Q9HNZ9	Q9hnz9 halobacteri
24	50	34.5	516	10	Q9SSP5	Q9ssp5 arabidopsis
25	50	34.5	851	11	Q8K2H3	Q8k2h3 mus musculu
26	49.5	34.1	153	5	Q9UB33	Q9ub33 anopheles g
27	49.5	34.1	401	5	O97407	O97407 anopheles g
28	49.5	34.1	505	8	O47148	O47148 menziesia c
29	49.5	34.1	506	8	O47149	O47149 rhododendro
30	49.5	34.1	506	8	O47171	O47171 rhododendro
31	49.5	34.1	506	8	O63960	O63960 rhododendro
32	49.5	34.1	506	8	O62982	O62982 rhododendro
33	49.5	34.1	506	8	O62975	O62975 rhododendro
34	49.5	34.1	506	8	O62972	O62972 rhododendro
35	49.5	34.1	506	8	O62989	O62989 rhododendro
36	49.5	34.1	506	8	O62978	O62978 rhododendro
37	49.5	34.1	506	8	O47155	O47155 rhododendro
38	49.5	34.1	506	8	O47152	O47152 rhododendro
39	49.5	34.1	506	8	O47173	O47173 rhododendro
40	49.5	34.1	506	8	O62990	O62990 rhododendro
41	49.5	34.1	506	8	O62974	O62974 rhododendro
42	49.5	34.1	506	8	O62993	O62993 menziesia m
43	49.5	34.1	506	8	O47170	O47170 rhododendro
44	49.5	34.1	506	8	O47174	O47174 rhododendro
45	49.5	34.1	506	8	O62983	O62983 rhododendro

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 60.0%; Score 87; DB 13; Length 146;
Best Local Similarity 65.2%; Pred. No. 6.7e-05;
Matches 15; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          2 LWAAQRYGRELRRMSDEFEGSFK 24
              |||| :||::||||||||:   |
Db         89 LWAAKKYGQQLRRMSDEFDKGMK 111
```

RESULT 2
Q8K016
ID Q8K016 PRELIMINARY; PRT; 946 AA.
AC Q8K016;

```
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Inter-alpha trypsin inhibitor, heavy chain 2.
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC034341; AAH34341.1; -.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE   946 AA;  105945 MW;  8B17DBA71B85BC5C CRC64;

  Query Match             41.4%;  Score 60;  DB 11;  Length 946;
  Best Local Similarity   37.0%;  Pred. No. 4.9;
  Matches   10;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              |:|  :  |     : | ||| |:|:|
Db        212 NVWIMEPQGMRFLHVPDTFEGHFQGVP 238


RESULT 3
Q8VJS3
ID   Q8VJS3      PRELIMINARY;      PRT;   196 AA.
AC   Q8VJS3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   IS1607, transposase.
GN   MT2070.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE007058; AAK46348.1; -.
DR   TIGR; MT2070; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
SQ   SEQUENCE   196 AA;  21349 MW;  C145A8D836FD9C2D CRC64;

  Query Match             36.6%;  Score 53;  DB 16;  Length 196;
  Best Local Similarity   58.8%;  Pred. No. 9.1;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        134 NLWAADRYNRAIARGHD 150


RESULT 4
Q10843
ID   Q10843      PRELIMINARY;      PRT;   223 AA.
AC   Q10843;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein Rv2014.
GN   RV2014 OR MTCY39.03C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SIMILARITY: TO M.PARATUBERCULOSIS IS900.
DR   EMBL; Z74025; CAA98415.1; -.
DR   TubercuList; Rv2014; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   223 AA;  24132 MW;  70456750017FEF37 CRC64;

  Query Match             36.6%;  Score 53;  DB 16;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 11;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSD 17
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 5
Q8ZY71
ID   Q8ZY71      PRELIMINARY;      PRT;   471 AA.
AC   Q8ZY71;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein PAE0922.
GN   PAE0922.
OS   Pyrobaculum aerophilum.
OC   Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC   Thermoproteaceae; Pyrobaculum.
OX   NCBI_TaxID=13773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX   PubMed=11792869;
RA   Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA   Miller J.H.;
RT   "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT   aerophilum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR   EMBL; AE009793; AAL63125.1; -.
DR   InterPro; IPR006638; Elp3.
DR   InterPro; IPR000182; GCN5acetyltransf.
DR   Pfam; PF00583; Acetyltransf; 1.
DR   SMART; SM00729; Elp3; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   471 AA;  52952 MW;  3B1E36E8AEE2EF0A CRC64;
```

```
  Query Match             35.9%;  Score 52;  DB 17;  Length 471;
  Best Local Similarity   41.7%;  Pred. No. 34;
  Matches   10;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          3 WAAQRYGRELRRMSDEFEGSFKGL 26
              |     |||| |:::|  | |  |
Db        404 WQHSGMGRELMRLAEEIAGEFGAL 427


RESULT 6
Q9GLY6
ID   Q9GLY6      PRELIMINARY;      PRT;   946 AA.
AC   Q9GLY6;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain2.
GN   ITI-HC2.
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Suzuki A.;
RT   "Rabbit inter-alpha-trypsin inhibitor heavy chain.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB050592; BAB17301.1; -.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE   946 AA;  106240 MW;  B7AF05434B228CC5 CRC64;

  Query Match             35.9%;  Score 52;  DB 6;  Length 946;
  Best Local Similarity   33.3%;  Pred. No. 74;
  Matches    9;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          1 NLWAAQRYGRELRRMSDEFEGSFKGLP 27
              ::|  :  |     : | ||| | |:|
Db        212 DVWIIEPQGLRFLHVPDTFEGHFDGVP 238


RESULT 7
Q9SEG0
ID   Q9SEG0      PRELIMINARY;      PRT;   231 AA.
AC   Q9SEG0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Apetala3 (Fragment).
OS   Arabidopsis lyrata.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=59689;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=99404148; PubMed=10474900;
RA   Lawton-Rauh A.L., Buckler E.S. IV, Purugganan M.D.;
RT   "Patterns of molecular evolution among paralogous floral homeotic
RT   genes.";
RL   Mol. Biol. Evol. 16:1037-1045(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF143380; AAF25590.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
FT   NON_TER     231    231
SQ   SEQUENCE   231 AA;  27176 MW;  A67CAE1EEBD8F7AA CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 231;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |         :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 8
Q9SQ20
ID   Q9SQ20      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ20;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Corscalla;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115806; AAD51895.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27342 MW;  BDFDCB59B73F4601 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |         :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 9
Q9SQ22
ID   Q9SQ22      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ22;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
```

```
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Li-8;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115801; AAD51890.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27267 MW;  42A852D697E22A65 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 10
Q9SQ17
ID   Q9SQ17      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ17;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115809; AAD51898.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27284 MW;  04FCFC55B73C7729 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 11
Q9SQ19
ID   Q9SQ19      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ19;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Bretagny;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115807; AAD51896.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27311 MW;  71AE593FB8A67EC3 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 12
Q9SQ21
ID   Q9SQ21      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ21;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Kent;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115805; AAD51894.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27286 MW;  66976305B88B63E3 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 13
Q8LB79
ID   Q8LB79      PRELIMINARY;      PRT;   232 AA.
AC   Q8LB79;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Floral homeotic protein APETALA3 (AP3).
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA   Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT   "Full-length messenger RNA sequences greatly improve genome
RT   annotation.";
RL   Genome Biol. 0:0-0(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA   Feldmann K.;
RT   "Full-Length cDNA from Arabidopsis thaliana.";
RL   Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AY087369; AAM64919.1; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27339 MW;  CC90703F959CFAD5 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLNKLDIQELRRLEDEMENTFK 133


RESULT 14
Q9SQ16
ID   Q9SQ16      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ16;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Lisse;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115810; AAD51899.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27314 MW;  DB8CA1FC835557D6 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 15
Q9SQ15
ID   Q9SQ15      PRELIMINARY;      PRT;   232 AA.
AC   Q9SQ15;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Floral homeotic protein AP3.
GN   APETALA3.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
```

```
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Kas-1;
RX   MEDLINE=99126449; PubMed=9927474;
RA   Purugganan M.D., Suddith J.I.;
RT   "Molecular population genetics of floral homeotic loci. Departures
RT   from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT   genes of Arabidopsis thaliana.";
RL   Genetics 151:839-848(1999).
CC   -!- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC       FACTORS.
DR   EMBL; AF115812; AAD51901.1; -.
DR   HSSP; P11746; 1MNM.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ   SEQUENCE   232 AA;  27300 MW;  5CA05FD44F824DF0 CRC64;

  Query Match             35.2%;  Score 51;  DB 10;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 22;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          6 QRYG--------RELRRMSDEFEGSFK 24
              || |        :|||| : || | :||
Db        107 QRVGECLDELDIQELRRLEDEMENTFK 133


Search completed: September 15, 2003, 17:25:46
Job time : 30.3143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 39.6 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	148	100.0	28	21	AAB37055	Bcl2 polypeptide B
2	143	96.6	27	21	AAB37056	Bcl2 polypeptide B
3	138	93.2	26	21	AAB37001	Bcl2 polypeptide B
4	138	93.2	26	21	AAB37002	Bcl2 polypeptide B
5	138	93.2	27	21	AAB37003	Bcl2 polypeptide B
6	138	93.2	162	22	AAB70370	Shorter murine BAD
7	138	93.2	204	17	AAR95168	bcl-x(L)/bcl-2 ass
8	138	93.2	204	19	AAW61315	Murine BCL-XL/BCL-
9	138	93.2	204	19	AAW61316	Mutant BCL-XL/BCL-
10	138	93.2	204	19	AAW61317	Mutant BCL-XL/BCL-
11	138	93.2	204	19	AAW61318	Mutant BCL-XL/BCL-
12	138	93.2	204	19	AAW58832	Murine BAD protein
13	138	93.2	204	22	AAB70369	Longer murine BAD
14	138	93.2	204	24	ABR39082	Murine BAD protein
15	138	93.2	567	22	AAU00220	Bad-DTTR apoptosis
16	114	77.0	24	23	AAU78627	Human Bad peptide
17	114	77.0	25	23	ABP56161	PTPC-interacting T
18	114	77.0	25	23	ABG78484	Mutant Bcl2 compet
19	114	77.0	25	23	ABG78493	Mutant Bcl2 compet
20	114	77.0	25	23	AAU78610	Human Bad peptide
21	114	77.0	25	23	AAU78620	Human Bad peptide
22	114	77.0	166	18	AAW32476	BBC6 protein for r
23	114	77.0	168	19	AAW55779	Human Bcl-xL/Bcl-2
24	114	77.0	168	21	AAB13512	Human cell prolife
25	114	77.0	168	22	AAB70368	Human BAD mutant a
26	114	77.0	168	22	AAB48287	Human Bad protein.
27	114	77.0	168	22	AAG67688	Amino acid sequenc
28	114	77.0	168	24	ABR39081	Human BAD protein
29	114	77.0	201	23	ABP41630	Human ovarian anti
30	113	76.4	23	17	AAR95166	bcl-x(L)/bcl-2 ass
31	111	75.0	25	23	ABG78490	Mutant Bcl2 compet
32	111	75.0	25	23	AAU78617	Human Bad peptide
33	110	74.3	25	23	ABG78488	Mutant Bcl2 compet
34	110	74.3	25	23	ABG78489	Mutant Bcl2 compet
35	110	74.3	25	23	AAU78615	Human Bad peptide
36	110	74.3	25	23	AAU78616	Human Bad peptide
37	109	73.6	23	23	AAU78628	Human Bad peptide
38	109	73.6	25	23	ABG78486	Mutant Bcl2 compet
39	109	73.6	25	23	ABG78492	Mutant Bcl2 compet
40	109	73.6	25	23	ABG78497	Mutant Bcl2 compet
41	109	73.6	25	23	AAU78612	Human Bad peptide
42	109	73.6	25	23	AAU78619	Human Bad peptide
43	109	73.6	25	23	AAU78624	Human Bad peptide
44	108	73.0	25	23	ABG78485	Mutant Bcl2 compet
45	108	73.0	25	23	AAU78611	Human Bad peptide

ALIGNMENTS

RESULT 1
AAB37055
ID AAB37055 standard; peptide; 28 AA.
XX
AC AAB37055;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #55.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.

XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 28 AA;

Query Match 100.0%; Score 148; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 1.4e-15;
Matches 28; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28
||||||||||||||||||||||||||||
Db 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

RESULT 2
AAB37056
ID AAB37056 standard; peptide; 27 AA.
XX
AC AAB37056;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #56.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 96.6%; Score 143; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.1e-15;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27
|||||||||||||||||||||||||||
Db 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 3
AAB37001
ID AAB37001 standard; peptide; 26 AA.
XX
AC AAB37001;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #1.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.

XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 93.2%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.6e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 4
AAB37002
ID AAB37002 standard; peptide; 26 AA.
XX
AC AAB37002;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #2.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 93.2%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.6e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 5
AAB37003
ID AAB37003 standard; peptide; 27 AA.
XX
AC AAB37003;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #3.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.

XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 93.2%; Score 138; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 6
AAB70370
ID AAB70370 standard; protein; 162 AA.
XX
AC AAB70370;
XX
DT 02-MAY-2001 (first entry)
XX
DE Shorter murine BAD mutant amino acid sequence SEQ ID NO:3.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 7; Page 148-149; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed shorter murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 162 AA;

Query Match 93.2%; Score 138; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 98 NLWAAQRYGRELRRMSDEFEGSFKGL 123

RESULT 7
AAR95168
ID AAR95168 standard; Protein; 204 AA.
XX
AC AAR95168;
XX
DT 06-JAN-1997 (first entry)
XX
DE bcl-x(L)/bcl-2 associated death promoter protein.
XX
KW Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS Mus musculus.
XX

```
FH   Key             Location/Qualifiers
FT   Region          147..149
FT                   /note= "BH1 conserved amino acids"
FT   Region          191..192
FT                   /note= "BH2 conserved amino acids"
FT   Domain          38..61
FT                   /note= "PEST sequence"
FT   Domain          111..130
FT                   /note= "PEST sequence"
XX
PN   WO9613614-A1.
XX
PD   09-MAY-1996.
XX
PF   31-OCT-1995;    95WO-US14246.
XX
PR   31-OCT-1994;    94US-0333565.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1996-251465/25.
DR   N-PSDB; AAT29479.
XX
PT   Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT   useful to treat neoplasia and apoptosis and to identify agents
PT   inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS   Claim 3; Fig 1; 130pp; English.
XX
CC   This sequence represents the murine bcl-x(L)/bcl-2 associated death
CC   promoter (Bad) gene.  Bad is a 22.1 kD protein which interacts with
CC   bcl-2 and bcl-x proteins and regulates cell death.  It has homology
CC   to the bcl-2-related family clustered in the BH1 and BH2 domain.  Bad
CC   has been found to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid
CC   assays and in vivo in mammalian cells.  Overexpressed Bad counters the
CC   death inhibitory activity of bcl-x(L), but is much less effective at
CC   countering the death inhibitory activity of bcl-2.  Bad expression can
CC   accelerate apoptotic cell death induced by cytokine deprivation in an
CC   IL-3 dependent cell line expressing bcl-x(L), and its also counters the
CC   death repressor activity of bcl-x(L).  Bad competes with Bax for binding
CC   to bcl-x(L).  Bad may be used to identify agents which inhibit its
CC   binding to bcl-2 or bcl-x(L) to form heterodimers.  Such agents may be
CC   used to treat neurodegenerative diseases, immunodeficiency diseases,
CC   e.g. AIDS, senescence or ischaemia.
XX
SQ   Sequence   204 AA;

  Query Match             93.2%;  Score 138;  DB 17;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 8
AAW61315
ID   AAW61315 standard; Protein; 204 AA.
XX
AC   AAW61315;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Murine BCL-XL/BCL-2 associated cell death regulator.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;    97WO-US19175.
XX
PR   18-OCT-1996;    96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27833.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 1; Fig 10; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence is the murine BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             93.2%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 9
AAW61316
ID   AAW61316 standard; Protein; 204 AA.
XX
AC   AAW61316;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #1.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
```

```
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27834.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 59; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             93.2%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 10
AAW61317
ID   AAW61317 standard; Protein; 204 AA.
XX
AC   AAW61317;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #2.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27835.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 60; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             93.2%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 11
AAW61318
ID   AAW61318 standard; Protein; 204 AA.
XX
AC   AAW61318;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #3.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
```

XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27836.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 60-61; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 93.2%; Score 138; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 12
AAW58832
ID AAW58832 standard; protein; 204 AA.
XX
AC AAW58832;
XX
DT 23-JUL-1998 (first entry)
XX
DE Murine BAD protein.
XX
KW BAD protein; Bcl-XL/Bcl-2 associated cell death regulator; 14-3-3;
KW serine phosphorylation; post-translational modification; apoptosis;
KW signal transduction regulator; phosphoserine phosphatase; senescence;
KW immunodeficiency disease; neurodegenerative disease; infertility;
KW cancer, viral infection; lymphoproliferative condition; arthritis;
KW inflammation; autoimmune diseases.
XX
OS Mus sp.
XX
PN WO9809643-A1.
XX
PD 12-MAR-1998.
XX
PF 09-SEP-1997; 97WO-US15871.
XX
PR 09-SEP-1996; 96US-0707868.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-207049/18.
XX
PT Serine-phosphorylated Bcl-X-l/Bcl-2 Associated cell Death regulator
PT polypeptide - useful for modulation of apoptosis associated with,
PT e.g. cancer and immunodeficiency diseases
XX
PS Claim 3; Fig 8; 61pp; English.
XX
CC This sequence represents a novel serine-phosphorylated protein, BAD
CC (Bcl-Xl/Bcl-2 associated cell death regulator). The serine residue is
CC phosporylated in a post-translational modification and allows binding
CC to the 14-3-3 protein which is a signal transduction regulator.
CC Modulators of phosphorylated BAD, which act through inhibition/activation
CC of a phosphoserine phosphatase, are useful for preventing/treating
CC increased/decreased apoptosis in a cell. The increased apoptosis may
CC result from immunodeficiency diseases, senescence, neurodegenerative
CC disease, ischaemic cell death, reperfusion cell death, infertility and
CC wound-healing. Decreased apoptosis may result from cancer, viral
CC infection, lymphoproliferative conditions, arthritis, infertility,
CC inflammation and autoimmune diseases. Measuring the amount of
CC phosphorylated compared to unphosphorylated BAD polypeptide and/or total
CC BAD in a cell is useful for determining the apoptotic state of a cell.
XX
SQ Sequence 204 AA;

Query Match 93.2%; Score 138; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 13
AAB70369
ID AAB70369 standard; protein; 204 AA.
XX
AC AAB70369;
XX
DT 02-MAY-2001 (first entry)
XX
DE Longer murine BAD mutant amino acid sequence SEQ ID NO:2.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.

XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 4; Page 148; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed longer murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 204 AA;

Query Match 93.2%; Score 138; DB 22; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 14
ABR39082
ID ABR39082 standard; Protein; 204 AA.
XX
AC ABR39082;
XX
DT 10-MAY-2003 (first entry)
XX
DE Murine BAD protein SEQ ID NO:4.
XX
KW Murine; BAD; herpes simplex virus; HSV; US3; herpes virus; apoptosis;
KW virucide; infection.
XX
OS Mus musculus.
XX
PN WO2003012049-A2.
XX
PD 13-FEB-2003.
XX
PF 31-JUL-2002; 2002WO-US24177.
XX
PR 31-JUL-2001; 2001US-308929P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Munger J, Roizman B;
XX
DR WPI; 2003-248168/24.
DR N-PSDB; ABZ81201.
XX
PT Inducing apoptosis in a cell infected with herpes simplex virus, HSV,
PT by administering to the cell, a composition comprising an agent that
PT inhibits phosphorylation of pro-apoptotic polypeptide BAD by HSV US3 -
XX
PS Claim 15; Page 168; 192pp; English.
XX
CC The present invention describes a method (M1) for inducing apoptosis in
CC a cell infected with herpes simplex virus (HSV), which comprises
CC administering to the cell, a composition having an agent that inhibits
CC phosphorylation of pro-apoptotic polypeptide BAD by HSV US3. Also
CC described is a method (M2) for treating a patient infected with HSV, by
CC administering to the patient, a composition comprising a peptide
CC comprising a sequence of 4-100 continuous amino acids of a 168 residue
CC amino acid sequence (see ABR39081), where the peptide comprises ser112,
CC ser135, or ser155, or their combinations. BAD has virucide activity.
CC M1 is useful for inducing apoptosis in a cell infected with HSV, where
CC the cell is in a human. M2 is useful for treating a patient infected
CC with HSV. The present sequence represents murine BAD, which is used in
CC the exemplification of the present invention.
XX
SQ Sequence 204 AA;

Query Match 93.2%; Score 138; DB 24; Length 204;
Best Local Similarity 100.0%; Pred. No. 4.5e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 15
AAU00220
ID AAU00220 standard; Protein; 567 AA.
XX
AC AAU00220;
XX
DT 31-MAY-2001 (first entry)
XX
DE Bad-DTTR apoptosis-modifying fusion protein.
XX
KW Mouse; Bad-DTTR; apoptosis; cancer; spinal muscular atrophy;
KW diphtheria toxin receptor binding domain; DTR; neoplasm; tumour;
KW hyper-proliferation; Alzheimer's disease; neurodegenerative disorder;
KW transient ischaemic neuronal injury; stroke; spinal cord injury;
KW Huntington's disease.
XX
OS Chimeric - Mus sp.
OS Chimeric - Corynebacterium diptheriae.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Region 3..12
FT /note= "10x histidine tag"
XX
PN WO200112661-A2.
XX
PD 22-FEB-2001.
XX
PF 15-AUG-2000; 2000WO-US22293.
XX
PR 16-AUG-1999; 99US-0149220.
XX
PA (HARD) HARVARD COLLEGE.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Youle RJ, Liu X, Collier RJ;
XX
DR WPI; 2001-218343/22.
DR N-PSDB; AAS00248.
XX

PT Novel fusion protein for modifying apoptosis in target cell and
PT reducing apoptosis after transient ischaemic neuronal injury, has two
PT domains which targets protein to a cell and modifies apoptotic response
PT of cell -
XX
PS Claim 4; Page 59-61; 65pp; English.
XX
CC The sequence represents the amino acid sequence of Bad-DTTR apoptosis-
CC modifying fusion protein comprising Bad gene sequence fused via a short
CC linker to diphtheria toxin translocation domain (DTTR). The
CC functional apoptosis-modifying fusion protein is capable of binding a
CC target cell and integrating into or crossing a cellular membrane of the
CC target cell. The apoptosis-modifying fusion protein comprises at least
CC two domains: the DTR domain, which targets the fusion protein to the
CC target cell and the Bcl-XL domain, which modifies an apoptotic response
CC of the target cell. The fusion protein is useful for modifying
CC (inhibiting or enhancing) apoptosis in a target cell, such as neuron,
CC lymphocyte, cancer, neoplasm, macrophage, epithelial, stem, tumour or
CC hyper-proliferative cell or an adipocyte. It is also useful for reducing
CC apoptosis in a subject after transient ischaemic neuronal injury,
CC especially spinal cord injury. The fusion protein may be used to treat
CC various diseases and injury conditions through inhibition or enhancement
CC of apoptotic cellular response, including neurodegenerative disorders
CC such as Alzheimer's disease, Huntington's disease, spinal muscular
CC atrophy, stroke episodes and unregulated cell growth as in tumours and
CC various cancers. The apoptosis-modifying fusion protein can be delivered
CC effectively throughout the body and targeted to selective tissue and
CC cells.
XX
SQ Sequence 567 AA;

Query Match 93.2%; Score 138; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.4e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        161 NLWAAQRYGRELRRMSDEFEGSFKGL 186
```

Search completed: September 15, 2003, 17:22:14
Job time : 39.6 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 14.6 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	93.2	204	1	US-08-333-565-2	Sequence 2, Appli
2	138	93.2	204	2	US-08-661-479-2	Sequence 2, Appli
3	138	93.2	204	2	US-08-733-505A-1	Sequence 1, Appli
4	138	93.2	204	2	US-08-733-505A-12	Sequence 12, Appl
5	138	93.2	204	2	US-08-733-505A-13	Sequence 13, Appl
6	138	93.2	204	2	US-08-733-505A-14	Sequence 14, Appl
7	135	91.2	204	2	US-08-717-123-3	Sequence 3, Appli
8	135	91.2	204	4	US-09-375-257-3	Sequence 3, Appli
9	114	77.0	166	1	US-08-665-617-2	Sequence 2, Appli
10	114	77.0	168	2	US-08-717-123-2	Sequence 2, Appli
11	114	77.0	168	3	US-08-985-335-1	Sequence 1, Appli
12	114	77.0	168	3	US-08-985-335-7	Sequence 7, Appli
13	114	77.0	168	3	US-09-410-372-1	Sequence 1, Appli
14	114	77.0	168	3	US-09-410-372-7	Sequence 7, Appli
15	114	77.0	168	4	US-09-375-257-2	Sequence 2, Appli
16	113	76.4	23	1	US-08-333-565-10	Sequence 10, Appl
17	113	76.4	23	2	US-08-661-479-10	Sequence 10, Appl
18	102	68.9	59	2	US-08-733-505A-55	Sequence 55, Appl
19	102	68.9	59	2	US-08-733-505A-56	Sequence 56, Appl
20	102	68.9	59	2	US-08-733-505A-57	Sequence 57, Appl
21	102	68.9	59	2	US-08-733-505A-58	Sequence 58, Appl
22	86	58.1	16	1	US-08-333-565-26	Sequence 26, Appl
23	86	58.1	16	2	US-08-661-479-26	Sequence 26, Appl
24	61	41.2	11	2	US-08-733-505A-34	Sequence 34, Appl
25	61	41.2	11	2	US-08-706-741B-69	Sequence 69, Appl
26	61	41.2	11	2	US-08-924-695A-69	Sequence 69, Appl
27	51	34.5	66	2	US-08-867-087B-40	Sequence 40, Appl
28	48.5	32.8	904	4	US-09-328-352-4656	Sequence 4656, Ap
29	46	31.1	610	4	US-09-252-991A-19594	Sequence 19594, A
30	46	31.1	946	3	US-09-074-579-3	Sequence 3, Appli
31	46	31.1	946	3	US-09-388-774-3	Sequence 3, Appli
32	46	31.1	946	4	US-09-546-153-1	Sequence 1, Appli
33	45.5	30.7	906	4	US-09-252-991A-31458	Sequence 31458, A
34	45	30.4	229	4	US-09-252-991A-23807	Sequence 23807, A
35	45	30.4	303	4	US-09-328-352-5164	Sequence 5164, Ap
36	45	30.4	356	4	US-09-235-103-2	Sequence 2, Appli
37	45	30.4	356	4	US-09-235-103-4	Sequence 4, Appli
38	45	30.4	1064	4	US-09-252-991A-17508	Sequence 17508, A
39	44.5	30.1	903	4	US-09-252-991A-28775	Sequence 28775, A
40	44	29.7	125	4	US-09-328-352-7449	Sequence 7449, Ap
41	44	29.7	263	4	US-09-651-656-27	Sequence 27, Appl
42	44	29.7	263	4	US-09-650-855-27	Sequence 27, Appl
43	44	29.7	277	4	US-09-252-991A-28581	Sequence 28581, A
44	44	29.7	877	4	US-09-206-551-20	Sequence 20, Appl
45	44	29.7	1125	4	US-09-252-991A-18729	Sequence 18729, A

ALIGNMENTS

```
RESULT 1
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-333-565-2

  Query Match             93.2%;  Score 138;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 2
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             93.2%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 3
US-08-733-505A-1
; Sequence 1, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-733-505A-1

  Query Match             93.2%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 4
US-08-733-505A-12
; Sequence 12, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-12

  Query Match             93.2%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 5
US-08-733-505A-13
; Sequence 13, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-13

  Query Match             93.2%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 6
US-08-733-505A-14
; Sequence 14, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-14

  Query Match             93.2%;  Score 138;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 7
US-08-717-123-3
; Sequence 3, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
```

```
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-717-123-3

  Query Match            91.2%;  Score 135;  DB 2;  Length 204;
  Best Local Similarity  96.2%;  Pred. No. 4.2e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 8
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match            91.2%;  Score 135;  DB 4;  Length 204;
  Best Local Similarity  96.2%;  Pred. No. 4.2e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 9
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match            77.0%;  Score 114;  DB 1;  Length 166;
  Best Local Similarity  91.7%;  Pred. No. 5.4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||||  |||
Db        101 NLWAAQRYGRELRRMSDEFVDSFK 124


RESULT 10
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match            77.0%;  Score 114;  DB 2;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126
```

```
RESULT 11
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match             77.0%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 12
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match             77.0%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 13
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match             77.0%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 14
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match             77.0%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 15
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match             77.0%;  Score 114;  DB 4;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.5e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


Search completed: September 15, 2003, 17:45:06
Job time : 14.6 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 22.2 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	135	91.2	204	9	US-09-922-378-3	Sequence 3, Appli
2	135	91.2	204	14	US-10-066-179-3	Sequence 3, Appli
3	114	77.0	25	15	US-10-059-261-258	Sequence 258, App
4	114	77.0	168	9	US-09-922-378-2	Sequence 2, Appli
5	114	77.0	168	9	US-09-894-657-1	Sequence 1, Appli
6	114	77.0	168	9	US-09-894-657-7	Sequence 7, Appli
7	114	77.0	168	14	US-10-066-179-2	Sequence 2, Appli
8	71	48.0	15	15	US-10-174-105A-147	Sequence 147, App
9	52	35.1	215	15	US-10-156-761-9145	Sequence 9145, Ap
10	51	34.5	682	12	US-10-238-075-1077	Sequence 1077, Ap
11	47	31.8	35	15	US-10-092-750-1	Sequence 1, Appli
12	47	31.8	138	15	US-10-092-750-241	Sequence 241, App
13	46	31.1	946	9	US-09-828-423-3	Sequence 3, Appli
14	44	29.7	270	11	US-09-934-455-162	Sequence 162, App
15	44	29.7	272	15	US-10-156-761-11541	Sequence 11541, A
16	44	29.7	426	9	US-09-815-242-5704	Sequence 5704, Ap
17	44	29.7	699	14	US-10-008-355-8	Sequence 8, Appli
18	44	29.7	705	9	US-09-815-242-12463	Sequence 12463, A
19	44	29.7	712	14	US-10-008-355-2	Sequence 2, Appli
20	44	29.7	877	12	US-10-369-294-20	Sequence 20, Appl
21	43	29.1	94	9	US-09-864-761-39540	Sequence 39540, A
22	43	29.1	164	11	US-09-986-480-395	Sequence 395, App
23	43	29.1	213	9	US-09-843-846-18	Sequence 18, Appl
24	43	29.1	232	10	US-09-881-752A-238	Sequence 238, App
25	43	29.1	380	9	US-09-149-045-2	Sequence 2, Appli
26	43	29.1	380	15	US-10-166-359-2	Sequence 2, Appli
27	43	29.1	380	15	US-10-166-113-2	Sequence 2, Appli
28	43	29.1	380	15	US-10-166-357-2	Sequence 2, Appli
29	43	29.1	380	15	US-10-166-372-2	Sequence 2, Appli
30	43	29.1	380	15	US-10-184-722-3	Sequence 3, Appli
31	43	29.1	380	15	US-10-251-385-62	Sequence 62, Appl
32	43	29.1	380	15	US-10-251-385-198	Sequence 198, App
33	43	29.1	380	15	US-10-225-567A-233	Sequence 233, App
34	43	29.1	543	15	US-10-156-761-13485	Sequence 13485, A
35	43	29.1	571	9	US-09-815-242-11813	Sequence 11813, A
36	43	29.1	582	10	US-09-331-631A-22	Sequence 22, Appl
37	43	29.1	640	9	US-09-989-722-501	Sequence 501, App
38	43	29.1	640	9	US-09-989-723-501	Sequence 501, App
39	43	29.1	640	9	US-09-989-279-501	Sequence 501, App
40	43	29.1	640	9	US-09-989-727-501	Sequence 501, App
41	43	29.1	640	10	US-09-989-731-501	Sequence 501, App
42	43	29.1	640	10	US-09-989-732-501	Sequence 501, App
43	43	29.1	640	10	US-09-991-073-501	Sequence 501, App
44	43	29.1	640	10	US-09-909-320-292	Sequence 292, App
45	43	29.1	640	10	US-09-990-442-501	Sequence 501, App

ALIGNMENTS

RESULT 1
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE
; FILE REFERENCE: 480140.428D3
; CURRENT APPLICATION NUMBER: US/09/922,378
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-922-378-3

Query Match 91.2%; Score 135; DB 9; Length 204;
Best Local Similarity 96.2%; Pred. No. 8.2e-12;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:|||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165
```

RESULT 2
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE

```
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             91.2%;  Score 135;  DB 14;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 8.2e-12;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 3
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match             77.0%;  Score 114;  DB 15;  Length 25;
  Best Local Similarity   91.7%;  Pred. No. 1e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||  |||
Db          1 NLWAAQRYGRELRRMSDEFVDSFK 24


RESULT 4
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match             77.0%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 7.2e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 5
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: SYNORAB01
;            CLONE: 358673
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             77.0%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 7.2e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 6
US-09-894-657-7
```

```
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Yue, Henry
;                   Lal, Preeti
;                   Shah, Purvi
;                   Corley, Neil C.
;        TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                            PROLIFERATION
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/894,657
;             FILING DATE: 28-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/410,372
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0421 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 168 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GenBank
;             CLONE: 1683637
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match             77.0%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 7.2e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 7
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match             77.0%;  Score 114;  DB 14;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 7.2e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECIF
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHIN
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match             48.0%;  Score 71;  DB 15;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00097;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 GRELRRMSDEFEGS 23
              ||||||||||||||
Db          1 GRELRRMSDEFEGS 14


RESULT 9
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
```

```
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

  Query Match             35.1%;  Score 52;  DB 15;  Length 215;
  Best Local Similarity   45.5%;  Pred. No. 8.2;
  Matches   10;  Conservative    5;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          7 QRYGRELRRMSDEFEGSFKGLK 28
              :|:| :|||| || :|    |:
Db        108 ERWGGDLRRMRDEADGKVPELR 129


RESULT 10
US-10-238-075-1077
; Sequence 1077, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
;  APPLICANT: I.N.S.E.R.M.
;  TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are isolat
;  TITLE OF INVENTION:  E.coli, and biological uses of these polynucleotides and of thei
;  FILE REFERENCE: BLANDINE
;  CURRENT APPLICATION NUMBER: US/10/238,075
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 0003145
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 1576
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1077
;   LENGTH: 682
;   TYPE: PRT
;   ORGANISM: Escherichia coli
US-10-238-075-1077

  Query Match             34.5%;  Score 51;  DB 12;  Length 682;
  Best Local Similarity   40.7%;  Pred. No. 37;
  Matches   11;  Conservative    5;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27
              | :||||| | :: |: : |     ||
Db        608 KQIWAAQRNGAKVPRVRNGFTSMDIGL 634


RESULT 11
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-1

  Query Match             31.8%;  Score 47;  DB 15;  Length 35;
  Best Local Similarity   45.5%;  Pred. No. 6.7;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          3 LWAAQRYGRELRRMSDEFEGSF 24
              :| ||    ||||: |||   :
Db         15 IWIAQ----ELRRIGDEFNAYY 32


RESULT 12
US-10-092-750-241
; Sequence 241, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-241

  Query Match             31.8%;  Score 47;  DB 15;  Length 138;
  Best Local Similarity   45.5%;  Pred. No. 27;
  Matches   10;  Conservative    3;  Mismatches    5;  Indels    4;  Gaps    1;

Qy          3 LWAAQRYGRELRRMSDEFEGSF 24
              :| ||    ||||: |||   :
Db         86 IWIAQ----ELRRIGDEFNAYY 103


RESULT 13
US-09-828-423-3
; Sequence 3, Application US/09828423
; Patent No. US20020099178A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Guegler, Karl J.
;                   Patterson, Chandra
;        TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
;                            INHIBITOR HEAVY CHAIN PRECURSOR
;        NUMBER OF SEQUENCES: 5
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Drive
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/828,423
;             FILING DATE: 05-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/388,774
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Cerrone, Michael C
;             REGISTRATION NUMBER: 39,132
;             REFERENCE/DOCKET NUMBER: PF-0505 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
```

```
;             TELEX: <Unknown>
;   INFORMATION FOR SEQ ID NO: 3:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 946 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GENEBANK
;             CLONE: gi33985
;        SEQUENCE DESCRIPTION: SEQ ID NO: 3 :
US-09-828-423-3

  Query Match            31.1%;  Score 46;  DB 9;  Length 946;
  Best Local Similarity  30.8%;  Pred. No. 2.7e+02;
  Matches    8;  Conservative    5;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ::|  :  |     : | ||| | |:
Db        212 DVWVIEPQGLRFLHVPDTFEGHFDGV 237


RESULT 14
US-09-934-455-162
; Sequence 162, Application US/09934455
; Publication No. US20030121070A1
; GENERAL INFORMATION:
;  APPLICANT: Adam, Luc
;  APPLICANT:  Creelman, Robert
;  APPLICANT:  Dubell, Arnold
;  APPLICANT:  Heard, Jacqueline
;  APPLICANT:  Jiang, Cai-Zhong
;  APPLICANT:  Keddie, James
;  APPLICANT:  Pilgrim, Marsha
;  APPLICANT:  Ratcliffe, Oliver
;  APPLICANT:  Reuber, Lynne
;  APPLICANT:  Riechmann, Jose Luis
;  APPLICANT:  Yu, Guo-Liang
;  APPLICANT:  Pineda, Omaira
;  TITLE OF INVENTION: Genes for Modifying Plant Traits IV
;  FILE REFERENCE: MBI-0025
;  CURRENT APPLICATION NUMBER: US/09/934,455
;  CURRENT FILING DATE:  2001-08-22
;  PRIOR APPLICATION NUMBER: 60/227439
;  PRIOR FILING DATE: 2000-08-22
;  PRIOR APPLICATION NUMBER: MBI-0022
;  PRIOR FILING DATE: 2001-11-16
;  PRIOR APPLICATION NUMBER: MBI-0023
;  PRIOR FILING DATE: 2001-04-17
;  NUMBER OF SEQ ID NOS: 516
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 162
;   LENGTH: 270
;   TYPE: PRT
;   ORGANISM: Arabidopsis thaliana
US-09-934-455-162

  Query Match            29.7%;  Score 44;  DB 11;  Length 270;
  Best Local Similarity  40.7%;  Pred. No. 1.5e+02;
  Matches   11;  Conservative    4;  Mismatches    8;  Indels    4;  Gaps    1;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGLK 28
              |::  | |||: :| |     |  |||
Db         78 NMFQPQIYGRDFKRSS----SSMVGLK 100


RESULT 15
US-10-156-761-11541
; Sequence 11541, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11541
;   LENGTH: 272
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-11541

  Query Match            29.7%;  Score 44;  DB 15;  Length 272;
  Best Local Similarity  53.3%;  Pred. No. 1.5e+02;
  Matches    8;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          4 WAAQRYGRELRRMSD 18
              | |  :| |||| :|
Db         29 WIAAAHGAELRRAAD 43


Search completed: September 15, 2003, 17:47:53
Job time : 22.2 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 12.6 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	93.2	204	2	A55671	bad protein - mous
2	54	36.5	946	2	JC5575	inter-alpha-trypsi
3	53	35.8	223	2	D70760	hypothetical prote
4	53	35.8	946	2	S54354	inter-alpha-inhibi
5	52	35.1	370	2	S38185	2-dehydro-3-deoxy-
6	51	34.5	232	2	A42095	floral homeotic pr
7	50	33.8	374	2	C84338	spermidine/putresc
8	50	33.8	516	2	A96753	probable threonine
9	50	33.8	1378	2	A81393	DNA-directed RNA p
10	49.5	33.4	127	2	AI1210	glycerol-3-phospha
11	49	33.1	453	2	E83517	conserved hypothet
12	48.5	32.8	134	2	S40376	Ig kappa chain - h
13	48.5	32.8	314	2	T02975	annexin P35 - maiz
14	48	32.4	206	2	C36365	transforming prote
15	48	32.4	220	2	F72289	oxidoreductase, so
16	48	32.4	526	2	T08545	threonine synthase
17	48	32.4	1164	2	T24806	hypothetical prote
18	47.5	32.1	334	2	A39172	Antho-RFamide neur
19	47.5	32.1	1140	2	T09486	hypothetical prote
20	47	31.8	287	2	S43852	neuropeptide Pol-R
21	47	31.8	597	2	G82308	oxaloacetate decar
22	47	31.8	967	2	F82668	oxoglutarate dehyd
23	47	31.8	5138	2	B96695	hypothetical prote
24	46.5	31.4	314	2	T02961	annexin P33 - maiz
25	46.5	31.4	435	2	A44308	Antho-RFamide prec
26	46	31.1	165	2	S59899	chlorocruorin chai
27	46	31.1	399	2	T35440	probable polyamine
28	46	31.1	946	1	IYHU2	inter-alpha-trypsi
29	45.5	30.7	261	2	G69510	conserved hypothet
30	45.5	30.7	327	2	AF2859	conserved hypothet
31	45.5	30.7	327	2	D97636	probable secreted
32	45.5	30.7	562	2	C71473	hypothetical prote
33	45.5	30.7	905	2	G83314	NADH dehydrogenase
34	45.5	30.7	1014	2	T36031	excinuclease ABC c
35	45	30.4	273	2	S06736	photosystem II oxy
36	45	30.4	273	2	AG2287	manganese-stabilzi
37	45	30.4	295	2	F83201	conserved hypothet
38	45	30.4	346	2	H95406	conserved hypothet
39	45	30.4	486	2	T31294	hypothetical prote
40	45	30.4	591	2	B44465	sodium ion pump ox
41	45	30.4	591	2	AB0509	oxaloacetate decar
42	45	30.4	591	2	AE0909	oxaloacetate decar
43	45	30.4	596	2	A28088	oxaloacetate decar
44	45	30.4	715	2	S52675	probable membrane
45	45	30.4	864	1	VCLJG4	env polyprotein -

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 93.2%; Score 138; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165
```

RESULT 2
JC5575
inter-alpha-trypsin inhibitor heavy chain 2 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
C;Accession: JC5575; PC4485
R;Nakatani, T.; Suzuki, Y.; Yamamoto, T.; Sinohara, H.
J. Biochem. 122, 71-82, 1997
A;Title: Molecular cloning and sequencing of cDNAs encoding three heavy-chain precurs
sin inhibitor heavy chain family.
A;Reference number: JC5574; MUID:97420688; PMID:9276673
A;Accession: JC5575
A;Molecule type: mRNA
A;Residues: 1-946 <NAK>
A;Cross-references: DDBJ:D89286; NID:g1694689; PIDN:BAA13939.1; PID:g1694690
A;Experimental source: liver
A;Accession: PC4485
A;Molecule type: protein
A;Residues: 55-64;140-146;151-156;424-447;500-528;577-605 <NA2>
C;Comment: In the plasma three inter-alpha-trypsin inhibitor heavy chains 1, 2 and 3
that the complexes play important role for panceatic cancer.
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
F;261-264,717-916/Disulfide bonds: #status predicted

Query Match 36.5%; Score 54; DB 2; Length 946;
Best Local Similarity 34.6%; Pred. No. 11;

```
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGV 237


RESULT 3
D70760
hypothetical protein Rv2014 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70760
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70760
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <COL>
A;Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CAA98415.1; PID:e1299911;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2014

  Query Match             35.8%;  Score 53;  DB 2;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 3.6;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 4
S54354
inter-alpha-inhibitor H2 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999
C;Accession: S54354
R;Chan, P.; Risler, J.L.; Raguenez, G.; Salier, J.P.
Biochem. J. 306, 505-512, 1995
A;Title: The three heavy-chain precursors for the inter-alpha-inhibitor family in mouse:
A;Reference number: S54353; MUID:95194326; PMID:7534067
A;Accession: S54354
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-946 <CHA>
A;Cross-references: EMBL:X70392; NID:g695633; PIDN:CAA49842.1; PID:g695634
C;Superfamily: inter-alpha-trypsin inhibitor complex component II

  Query Match             35.8%;  Score 53;  DB 2;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 16;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237


RESULT 5
S38185
2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) ARO4 - yeast (Saccharomyces ce
N;Alternate names: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DAHP synthase;
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C;Accession: S38185; S46126; S46130; JN0322; B48651
R;Doignon, F.; Biteau, N.; Aigle, M.; Crouzet, M.
Yeast 9, 1131-1137, 1993
A;Title: The complete sequence of a 6794 bp segment located on the right arm of chrom
A;Reference number: S38185; MUID:94078675; PMID:8256522
A;Accession: S38185
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <DOI>
A;Cross-references: GB:L20296; NID:g311101; PIDN:AAA65607.1; PID:g311102
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46126
A;Molecule type: DNA
A;Residues: 1-370 <ALJ>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46130
A;Molecule type: DNA
A;Residues: 1-370 <AIG>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Kuenzler, M.; Paravicini, G.; Egli, C.M.; Irniger, S.; Braus, G.H.
Gene 113, 67-74, 1992
A;Title: Cloning, primary structure and regulation of the ARO4 gene, encoding the tyr
A;Reference number: JN0322; MUID:92225349; PMID:1348717
A;Accession: JN0322
A;Molecule type: DNA
A;Residues: 1-204,208-370 <KUE>
A;Cross-references: EMBL:X61107
R;Kuenzler, M.; Balmelli, T.; Egli, C.M.; Paravicini, G.; Braus, G.H.
J. Bacteriol. 175, 5548-5558, 1993
A;Title: Cloning, primary structure, and regulation of the HIS7 gene encoding a bifur
A;Reference number: A48651; MUID:93374850; PMID:8366040
A;Accession: B48651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 352-370 <KU2>
A;Cross-references: GB:X61107
C;Comment: This enzyme catalyzes the condensation of phosphoenolpyruvate and D-erythr
C;Genetics:
A;Gene: SGD:ARO4
A;Cross-references: SGD:S0000453; MIPS:YBR249c
A;Map position: 2R
C;Function:
A;Description: aldehyde-lyase; carbon-carbon lyase
A;Pathway: aromatic amino acid biosynthesis; shikimate pathway
A;Note: first step in shikimate pathway
C;Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase
C;Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; cy

  Query Match             35.1%;  Score 52;  DB 2;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 8.4;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEG 22
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 6
A42095
floral homeotic protein APETALA3 (AP3) - Arabidopsis thaliana
N;Alternate names: homeotic protein APETALA3; MADS-box regulatory protein AP3
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A42095; S52633; T47593
R;Jack, T.; Brockman, L.L.; Meyerowitz, E.M.
Cell 68, 683-697, 1992
A;Title: The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is
A;Reference number: A42095; MUID:92154682; PMID:1346756
A;Accession: A42095
A;Status: preliminary
A;Molecule type: mRNA
```

A;Residues: 1-232 <JAC>
A;Cross-references: GB:M86357; NID:g166607; PIDN:AAA32740.1; PID:g166608
A;Experimental source: petals, stamens
A;Note: sequence extracted from NCBI backbone (NCBIN:82520, NCBIP:82521)
R;Okamoto, H.; Yano, A.; Shiraishi, H.; Okada, K.; Shimura, Y.
Plant Mol. Biol. 26, 465-472, 1994
A;Title: Genetic complementation of a floral homeotic mutation, apetala3, with an Arabid
A;Reference number: S52633; MUID:95036018; PMID:7948893
A;Accession: S52633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <OKA>
A;Cross-references: GB:D21125
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24469
A;Accession: T47593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <BLO>
A;Cross-references: EMBL:AL132971
A;Experimental source: cultivar Columbia; BAC clone T12E18
C;Genetics:
A;Map position: 3
A;Introns: 63/2; 85/3; 106/2; 139/3; 153/3; 168/3
A;Note: T12E18.30
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

Query Match 34.5%; Score 51; DB 2; Length 232;
Best Local Similarity 44.4%; Pred. No. 7.3;
Matches 12; Conservative 3; Mismatches 4; Indels 8; Gaps 1;

Qy 7 QRYG--------RELRRMSDEFEGSFK 25
|| | :|||: || | :||
Db 107 QRLGECLDELDIQELRRLEDEMENTFK 133

RESULT 7
C84338
spermidine/putrescine ABC transporter [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84338
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84338
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <STO>
A;Cross-references: GB:AE004437; NID:g10581314; PIDN:AAG20071.1; GSPDB:GN00138
C;Genetics:
A;Gene: potA2

Query Match 33.8%; Score 50; DB 2; Length 374;
Best Local Similarity 76.9%; Pred. No. 17;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 12 ELRRMSDEFEGSF 24
|||| :|| ||||
Db 197 ELRRLSDAVEGSF 209

RESULT 8
A96753
probable threonine synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96753
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alor
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE005173; NID:g5903070; PIDN:AAD55628.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3N23.1
A;Map position: 1

Query Match 33.8%; Score 50; DB 2; Length 516;
Best Local Similarity 35.3%; Pred. No. 23;
Matches 12; Conservative 7; Mismatches 7; Indels 8; Gaps 1;

Qy 2 NLWAAQRYGRELRRMSD--------EFEGSFKGL 27
||: |:|:|:: :|:| |||| |
Db 163 NLFWAERFGKQYLQMNDLWVKHCGISHTGSFKDL 196

RESULT 9
A81393
DNA-directed RNA polymerase (EC 2.7.7.6) beta chain Cj0478 [imported] - Campylobacte
C;Species: Campylobacter jejuni
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 03-Jun-2002
C;Accession: A81393
R;Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chi
C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Ba
Nature 403, 665-668, 2000
A;Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals
A;Reference number: A81250; MUID:20150912; PMID:10688204
A;Accession: A81393
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1378 <PAR>
A;Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB75116.1; PID:g69
A;Experimental source: serotype O2, strain NCTC 11168
C;Genetics:
A;Gene: rpoB; Cj0478
C;Superfamily: DNA-directed RNA polymerase beta chain
C;Keywords: nucleotidyltransferase

Query Match 33.8%; Score 50; DB 2; Length 1378;
Best Local Similarity 40.6%; Pred. No. 63;
Matches 13; Conservative 3; Mismatches 10; Indels 6; Gaps 2;

Qy 3 LWAAQRYG--RELRRM----SDEFEGSFKGLK 28
:|| : || || | ||: || | |
Db 1306 VWALEAYGAAHTLREMLTIKSDDVEGRFSAYK 1337

RESULT 10
AI1210
glycerol-3-phosphate cytidylyltransferase (gct), CDP-glycerol pyrophosphorylase (tei
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 01-Mar-2002
C;Accession: AI1210
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloe
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.

```
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1210
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-127 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAC99167.1; PID:g16410491; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: tagD
C;Superfamily: Bacillus subtilis glycerol-3-phosphate cytidylyltransferase tagD

  Query Match             33.4%;  Score 49.5;  DB 2;  Length 127;
  Best Local Similarity   36.4%;  Pred. No. 6.5;
  Matches   12;  Conservative    6;  Mismatches   10;  Indels    5;  Gaps    1;

Qy          1 KNLWAAQR-----YGRELRRMSDEFEGSFKGLK 28
              :| |  :|     || ::  | |::|| |  ||
Db         71 ENNWEQKRDDIEKYGIDVFVMGDDWEGEFDFLK 103


RESULT 11
E83517
conserved hypothetical protein PA1031 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83517
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <STO>
A;Cross-references: GB:AE004535; GB:AE004091; NID:g9946936; PIDN:AAG04420.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1031

  Query Match             33.1%;  Score 49;  DB 2;  Length 453;
  Best Local Similarity   55.6%;  Pred. No. 28;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          4 WAAQRYGR--ELRRMSDE 19
              ||::| ||  ||||:: |
Db         65 WASERQGREEELRRLASE 82


RESULT 12
S40376
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40376
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40376
A;Status: preliminary; translation not shown
A;Molecule type: mRNA
A;Residues: 1-134 <KLE>
A;Cross-references: EMBL:X72486; NID:g441440; PIDN:CAA51154.1; PID:g441441
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-113/Domain: immunoglobulin homology <IMM>

  Query Match             32.8%;  Score 48.5;  DB 2;  Length 134;
  Best Local Similarity   38.2%;  Pred. No. 9.7;
  Matches   13;  Conservative    1;  Mismatches    9;  Indels   11;  Gaps    1;

Qy          4 WAAQRYGRELRRM----------SDEFEGSFKG 26
              |  || ||  ||:           || | ||  |
Db         58 WFRQRPGRSPRRLIYNVSKRDSGVSDRFSGSGSG 91


RESULT 13
T02975
annexin P35 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02975
R;Battey, N.H.; James, N.C.; Greenland, A.J.
Plant Physiol. 112, 1391-1396, 1996
A;Title: CDNA isolation and gene expression of maize annexins P33 and P35.
A;Reference number: Z14796; MUID:97092863; PMID:8938425
A;Accession: T02975
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-314 <BAT>
A;Cross-references: EMBL:X98245; NID:g1370602; PIDN:CAA66901.1; PID:g1370603
A;Experimental source: cultivar clipper; root tip
C;Superfamily: annexin I; annexin repeat homology
F;14-85/Domain: annexin repeat homology <AXR>

  Query Match             32.8%;  Score 48.5;  DB 2;  Length 314;
  Best Local Similarity   47.6%;  Pred. No. 23;
  Matches   10;  Conservative    4;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          6 AQRYGRE-LRRMSDEFEGSFK 25
              |: ||:| || : ||  | |:
Db         54 AEAYGKELLRALGDEIHGKFE 74


RESULT 14
C36365
transforming protein homolog MRAS3 - Rhizomucor racemosus
C;Species: Rhizomucor racemosus
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 19-Jan-2001
C;Accession: C36365
R;Casale, W.L.; McConnell, D.G.; Wang, S.Y.; Lee, Y.J.; Linz, J.E.
Mol. Cell. Biol. 10, 6654-6663, 1990
A;Title: Expression of a gene family in the dimorphic fungus Mucor racemosus which ex
A;Reference number: A36365; MUID:91061774; PMID:1701021
A;Accession: C36365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <CAS>
A;Cross-references: GB:M55177
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;11-126/Domain: translation elongation factor Tu homology <ETU>
F;17-24/Region: nucleotide-binding motif A (P-loop)
F;123-126/Region: GTP-binding NKXD motif
F;153-155/Region: GTP-binding SAK/L motif
F;23,24,42,123,124,126,153/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #

  Query Match             32.4%;  Score 48;  DB 2;  Length 206;
  Best Local Similarity   62.5%;  Pred. No. 18;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         11 RELRRMSDEFEGSFKG 26
              ||:|||: | ||  ||
Db        169 REIRRMNKEQEGRSKG 184


RESULT 15
F72289
oxidoreductase, sol/devB family - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72289
```

R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <ARN>
A;Cross-references: GB:AE001772; GB:AE000512; NID:g4981693; PIDN:AAD36230.1; PID:g498170
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1154
C;Superfamily: yeast SOL3 protein

Query Match 32.4%; Score 48; DB 2; Length 220;
Best Local Similarity 34.8%; Pred. No. 19;
Matches 8; Conservative 8; Mismatches 7; Indels 0; Gaps 0;

```
Qy          5 AAQRYGRELRRMSDEFEGSFKGL 27
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133
```

Search completed: September 15, 2003, 17:27:03
Job time : 12.6 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 6.6 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	93.2	204	1	BAD_MOUSE	Q61337 mus musculu
2	138	93.2	205	1	BAD_RAT	O35147 rattus norv
3	114	77.0	168	1	BAD_HUMAN	Q92934 homo sapien
4	54	36.5	946	1	ITH2_MESAU	P97279 mesocricetu
5	53	35.8	946	1	ITH2_MOUSE	Q61703 mus musculu
6	52.5	35.5	506	1	MATK_LEDPA	O62992 ledum palus
7	52.5	35.5	506	1	MATK_RHOFR	O62984 rhododendro
8	52.5	35.5	506	1	MATK_RHOTS	O62991 rhododendro
9	52	35.1	370	1	AROG_YEAST	P32449 saccharomyc
10	51	34.5	232	1	AP3_ARATH	P35632 arabidopsis
11	51	34.5	851	1	CE05_MOUSE	Q8k2h3 mus musculu
12	50	33.8	1378	1	RPOB_CAMJE	Q46124 campylobact
13	49	33.1	453	1	RMUC_PSEAE	Q9i4u3 pseudomonas
14	48	32.4	205	1	RAS3_RHIRA	P22280 rhizomucor
15	48	32.4	220	1	6PGL_THEMA	Q9x0n8 thermotoga
16	48	32.4	519	1	THRC_SOLTU	Q9mt28 solanum tub
17	48	32.4	526	1	THRC_ARATH	Q9s7b5 arabidopsis
18	47.5	32.1	334	1	FMRA_CALPA	Q01133 calliactis
19	47.5	32.1	507	1	MATK_LOIPR	O47169 loiseleuria
20	47	31.8	198	1	BIM_HUMAN	O43521 homo sapien
21	47	31.8	287	1	PRFA_POLPE	P21259 polyorchis
22	46.5	31.4	429	1	FMR2_ANTEL	Q16994 anthopleura
23	46.5	31.4	435	1	FMR1_ANTEL	P10419 anthopleura
24	46	31.1	946	1	ITH2_HUMAN	P19823 homo sapien
25	45.5	30.7	1014	1	UVRA_STRCO	Q9z507 streptomyce
26	45	30.4	273	1	PSBO_ANASP	P13907 anabaena sp
27	45	30.4	328	1	SNF4_KLULA	Q9p869 kluyveromyc
28	45	30.4	590	1	DCOA_SALTY	Q03030 salmonella
29	45	30.4	595	1	DCOA_KLEPN	P13187 klebsiella
30	45	30.4	653	1	HT2A_HUMAN	Q13049 homo sapien
31	45	30.4	865	1	ENV_SIVAT	P05886 simian immu
32	45	30.4	915	1	CE05_HUMAN	Q9nyf5 homo sapien
33	45	30.4	1535	1	LML1_CAEEL	Q18823 caenorhabdi
34	45	30.4	5596	1	MDN1_HUMAN	Q9nu22 homo sapien
35	44.5	30.1	506	1	MATK_GAUPR	Q95gj0 gaultheria
36	44.5	30.1	512	1	MATK_LILTS	Q9gig3 lilium tsin
37	44.5	30.1	907	1	NUOG_ECOLI	P33602 escherichia
38	44.5	30.1	907	1	NUOG_SALTY	P33900 salmonella
39	44	29.7	196	1	BIM_MOUSE	O54918 mus musculu
40	44	29.7	196	1	BIM_RAT	O88498 rattus norv
41	44	29.7	262	1	END8_ECO57	Q8x9c6 escherichia
42	44	29.7	262	1	END8_ECOLI	P50465 escherichia
43	44	29.7	262	1	END8_SALTI	Q8z8d2 salmonella
44	44	29.7	262	1	END8_SALTY	Q8zqu6 salmonella
45	44	29.7	629	1	SYM_THEMA	O33925 thermotoga

ALIGNMENTS

RESULT 1
BAD_MOUSE
ID BAD_MOUSE STANDARD; PRT; 204 AA.
AC Q61337;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE 6) (Bcl-xL/Bcl-2 associated death promoter).
GN BAD OR BBC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Thymus;
RX MEDLINE=95136361; PubMed=7834748;
RA Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT promotes cell death.";
RL Cell 80:285-291(1995).
RN [2]
RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX MEDLINE=98022383; PubMed=9381178;
RA Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT "Interleukin-3-induced phosphorylation of BAD through the protein
RT kinase Akt.";
RL Science 278:687-689(1997).
RN [3]
RP MUTAGENESIS OF SERINE RESIDUES.
RX MEDLINE=20403302; PubMed=10949026;
RA Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA Greenberg M.E.;
RT "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT BH3 domain phosphorylation.";
RL Mol. Cell 6:41-51(2000).
CC -!- FUNCTION: Promotes cell death. Successfully competes for the
CC binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC of heterodimerization of these proteins with BAX. Can reverse the
CC death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC Appears to act as a link between growth factor receptor signaling
CC and the apoptotic pathways.
CC -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC phosphorylation, locates to the cytoplasm.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC with 14-3-3 proteins. This interaction then facilitates the
CC phosphorylation at Ser-155, a site within the BH3 domain, leading
CC to the release of Bcl-x(L) and the promotion of cell survival.

```
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             93.2%;  Score 138;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 2
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;
```

```
  Query Match             93.2%;  Score 138;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 2.5e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        141 NLWAAQRYGRELRRMSDEFEGSFKGL 166


RESULT 3
BAD_HUMAN
ID   BAD_HUMAN      STANDARD;      PRT;   168 AA.
AC   Q92934; O14803;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE   XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN   BAD OR BBC6 OR BCL2L8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT   "A human protein that interacts with Bcl-2 and have homology to mouse
RT   BAD.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX   MEDLINE=97083574; PubMed=8929532;
RA   Wang H.-G., Rapp U.R., Reed J.C.;
RT   "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL   Cell 87:629-638(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Takayama S., Reed J.C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND DIMERIZATION.
RC   TISSUE=Bone marrow;
RX   MEDLINE=98049554; PubMed=9388232;
RA   Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA   Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT   "Dimerization properties of human BAD.";
RL   J. Biol. Chem. 272:30866-30872(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match             77.0%;  Score 114;  DB 1;  Length 168;
```

```
  Best Local Similarity   91.7%;  Pred. No. 7.6e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 4
ITH2_MESAU
ID   ITH2_MESAU     STANDARD;      PRT;   946 AA.
AC   P97279;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2) (HC2).
GN   ITIH2.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97420688; PubMed=9276673;
RA   Nakatani T., Suzuki Y., Yamamoto T., Sinohara H.;
RT   "Molecular cloning and sequencing of cDNAs encoding three heavy-chain
RT   precursors of the inter-alpha-trypsin inhibitor in Syrian hamster:
RT   implications for the evolution of the inter-alpha-trypsin inhibitor
RT   heavy chain family.";
RL   J. Biochem. 122:71-82(1997).
RN   [2]
RP   SEQUENCE OF 55-64; 140-146; 151-156; 424-447; 500-528 AND 577-605,
RP   AND SUBUNITS.
RC   TISSUE=Plasma;
RX   MEDLINE=97018241; PubMed=8864857;
RA   Yamamoto T., Yamamoto K., Sinohara H.;
RT   "Inter-alpha-trypsin inhibitor and its related proteins in Syrian
RT   hamster urine and plasma.";
RL   J. Biochem. 120:145-152(1996).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89286; BAA13939.1; -.
DR   PIR; JC5575; JC5575.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                 H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                 (BY SIMILARITY).
FT   CONFLICT    510    510       V -> Y (IN REF. 2).
FT   CONFLICT    595    595       E -> I (IN REF. 2).
SQ   SEQUENCE   946 AA;  106580 MW;  CA8BF565458E7B2E CRC64;

  Query Match             36.5%;  Score 54;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 4.3;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGV 237


RESULT 5
ITH2_MOUSE
ID   ITH2_MOUSE     STANDARD;      PRT;   946 AA.
AC   Q61703;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2).
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6N; TISSUE=Liver;
RX   MEDLINE=95194326; PubMed=7534067;
RA   Chan P., Risler J.-L., Raguenez G., Salier J.-P.;
RT   "The three heavy-chain precursors for the inter-alpha-inhibitor
RT   family in mouse: new members of the multicopper oxidase protein group
RT   with differential transcription in liver and brain.";
RL   Biochem. J. 306:505-512(1995).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH LIVER AND BRAIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X70392; CAA49842.1; -.
DR   PIR; S54354; S54354.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                (BY SIMILARITY).
SQ   SEQUENCE   946 AA;  105927 MW;  40DB6716433ED9DC CRC64;

  Query Match             35.8%;  Score 53;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 6;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237


RESULT 6
MATK_LEDPA
ID   MATK_LEDPA     STANDARD;      PRT;   506 AA.
AC   O62992;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Ledum palustre (Wild rosemary).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75583;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012751; BAA25872.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60412 MW;  CFEA926307DAC85E CRC64;

  Query Match             35.5%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 3.6;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
                | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 7
MATK_RHOFR
ID   MATK_RHOFR     STANDARD;      PRT;   506 AA.
AC   O62984;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron ferrugineum (Alpenrose).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49622;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012741; BAA25862.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60534 MW;  ADA44B25E92436E8 CRC64;

  Query Match             35.5%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 3.6;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
                | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 8
MATK_RHOTS
ID   MATK_RHOTS     STANDARD;      PRT;   506 AA.
AC   O62991;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron tsusiophyllum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49629;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012750; BAA25871.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60569 MW;  AEE12FF8809C223E CRC64;

  Query Match           35.5%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity  37.5%;  Pred. No. 3.6;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 9
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match           35.1%;  Score 52;  DB 1;  Length 370;
  Best Local Similarity  47.6%;  Pred. No. 3;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEG 22
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 10
AP3_ARATH
ID   AP3_ARATH      STANDARD;      PRT;   232 AA.
AC   P35632; Q39003; Q8LB79; Q9S7Q3; Q9SQ14; Q9SQ15; Q9SQ16; Q9SQ17;
AC   Q9SQ18; Q9SQ19; Q9SQ20; Q9SQ21; Q9SQ22; Q9SX13;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Floral homeotic protein APETALA3.
GN   AP3 OR AT3G54340 OR T12E18_30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Petal;
RX MEDLINE=92154682; PubMed=1346756;
RA Jack T., Brockman L.L., Meyerowitz E.M.;
RT "The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS
RT box and is expressed in petals and stamens.";
RL Cell 68:683-697(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RX MEDLINE=95036018; PubMed=7948893;
RA Okamoto H., Yano A., Shiraishi H., Okada K., Shimura Y.;
RT "Genetic complementation of a floral homeotic mutation, apetala3,
RT with an Arabidopsis thaliana gene homologous to DEFICIENS of
RT Antirrhinum majus.";
RL Plant Mol. Biol. 26:465-472(1994).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC STRAIN=cv. Bla-1, cv. Bretagny, cv. Bs-1, cv. Bu-0, cv. Bu-2,
RC cv. Chi-1, cv. Co-1, cv. Columbia, cv. Corsacalla-1, cv. Cvi-0,
RC cv. Gr-3, cv. Jl-1, cv. Kas-1, cv. Kent, cv. Landsberg erecta,
RC cv. Li-3, cv. Li-8, and cv. Lisse;
RX MEDLINE=99126449; PubMed=9927474;
RA Purugganan M.D., Suddith J.I.;
RT "Molecular population genetics of floral homeotic loci: departures
RT from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT genes of Arabidopsis thaliana.";
RL Genetics 151:839-848(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [5]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP SEQUENCE OF 36-128 FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99311297; PubMed=10382288;
RA Brunel D., Froger N., Pelletier G.;
RT "Development of amplified consensus genetic markers (ACGM) in Brassica
RT napus from Arabidopsis thaliana sequences of known biological
RT function.";
RL Genome 42:387-402(1999).
RN [8]
RP FUNCTION.
RX PubMed=8565821;
RA Krizek B.A., Meyerowitz E.M.;
RT "The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient
RT to provide the B class organ identity function.";
RL Development 122:11-22(1996).
RN [9]
RP CHARACTERIZATION.
RX PubMed=8643482;
RA Riechmann J.L., Krizek B.A., Meyerowitz E.M.;
RT "Dimerization specificity of Arabidopsis MADS domain homeotic proteins
RT APETALA1, APETALA3, PISTILLATA, and AGAMOUS.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4793-4798(1996).
RN [10]
RP GENETIC REGULATION.
RX PubMed=11283333;
RA Ng M., Yanofsky M.F.;
RT "Activation of the Arabidopsis B class homeotic genes by APETALA1.";
RL Plant Cell 13:739-753(2001).
RN [11]
RP CHARACTERIZATION.
RX PubMed=11206550;
RA Honma T., Goto K.;
RT "Complexes of MADS-box proteins are sufficient to convert leaves into
RT floral organs.";
RL Nature 409:525-529(2001).
CC -!- FUNCTION: Probable transcription factor involved in the genetic
CC control of flower development. Is required for normal development
CC of petals and stamens in the wild-type flower. Forms an
CC heterodimer with PISTILLATA that is required for autoregulation of
CC both AP3 and PI genes. AP3/PI heterodimer interacts with APETALA1
CC or SEPALLATA3 to form a ternary complex that could be responsible
CC for the regulation of the genes involved in the flower
CC development.
CC -!- SUBUNIT: Forms an heterodimer with PISTILLATA, capable of binding
CC to CArG-box sequences. AP3/PI heterodimer binds AP1 or SEP3 to
CC form complexes.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in petals and stamens.
CC -!- INDUCTION: Positively regulated by the meristem identity proteins
CC APETALA1 and LEAFY with the cooperation of UFO.
CC -!- MISCELLANEOUS: Mutations in AP3 cause transformation of petals
CC into sepals and stamina into carpels.
CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS.
CC -!- SIMILARITY: Contains 1 K-box dimerization domain.
CC --

CC --
DR EMBL; M86357; AAA32740.1; -.
DR EMBL; D21125; BAA04665.1; -.
DR EMBL; AF115798; AAD51887.1; -.
DR EMBL; AF115799; AAD51888.1; -.

```
DR   EMBL; AF115800; AAD51889.1; -.
DR   EMBL; AF115801; AAD51890.1; -.
DR   EMBL; AF115802; AAD51891.1; -.
DR   EMBL; AF115803; AAD51892.1; -.
DR   EMBL; AF115804; AAD51893.1; -.
DR   EMBL; AF115805; AAD51894.1; -.
DR   EMBL; AF115806; AAD51895.1; -.
DR   EMBL; AF115807; AAD51896.1; -.
DR   EMBL; AF115808; AAD51897.1; -.
DR   EMBL; AF115809; AAD51898.1; -.
DR   EMBL; AF115810; AAD51899.1; -.
DR   EMBL; AF115811; AAD51900.1; -.
DR   EMBL; AF115812; AAD51901.1; -.
DR   EMBL; AF115813; AAD51902.1; -.
DR   EMBL; AF115814; AAD51903.1; -.
DR   EMBL; AL132971; CAB81799.1; -.
DR   EMBL; AY087369; AAM64919.1; -.
DR   EMBL; AY070397; AAL49893.1; -.
DR   EMBL; AY142590; AAN13159.1; -.
DR   EMBL; AF056541; AAD41557.1; -.
DR   PIR; A42095; A42095.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T01776; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Flowering; Transcription regulation; Activator; Developmental protein;
KW   Nuclear protein; DNA-binding; Coiled coil; Polymorphism.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       93    165       K-BOX.
FT   DOMAIN       75    164       COILED COIL (POTENTIAL).
FT   VARIANT      31     31       K -> R (in strain cv. Lisse).
FT   VARIANT      47     47       M -> T (in strain cv. Bretagny).
FT   VARIANT      61     61       N -> D (in strain cv. Corsacalla-1).
FT   VARIANT      73     73       T -> S (in strain cv. Li-8).
FT   VARIANT     109    109       L -> V (in strain cv. Kas-1).
FT   VARIANT     115    115       E -> K (in strains cv. Chi-1 and cv.
FT                                Gr-3).

  Query Match             34.5%;  Score 51;  DB 1;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 2.6;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          7 QRYG--------RELRRMSDEFEGSFK 25
              || |         :|||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 11
CE05_MOUSE
ID   CE05_MOUSE     STANDARD;      PRT;   851 AA.
AC   Q8K2H3;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Protein C5orf5 homolog.
GN   C5ORF5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Belongs to the FAM13 family.
CC   -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; BC031465; AAH31465.1; -.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   GTPase activation.
FT   DOMAIN       23    212       RHO-GAP.
FT   DOMAIN      189    256       GLU-RICH.
SQ   SEQUENCE   851 AA;  97054 MW;  C2B26669FB6DB2CE CRC64;

  Query Match             34.5%;  Score 51;  DB 1;  Length 851;
  Best Local Similarity   41.7%;  Pred. No. 11;
  Matches   10;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 KNLWAAQRYGRELRRMSDEFEGSF 24
              : || |:   ::||:|  ||| :|
Db        782 EQLWKARAEKKKLRKMLREFEEAF 805


RESULT 12
RPOB_CAMJE
ID   RPOB_CAMJE     STANDARD;      PRT;  1378 AA.
AC   Q46124; Q9PI31;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
DE   beta chain) (RNA polymerase beta subunit).
GN   RPOB OR CJ0478.
OS   Campylobacter jejuni.
OC   Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC   Campylobacteraceae; Campylobacter.
OX   NCBI_TaxID=197;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCTC 11168;
RX   MEDLINE=20150912; PubMed=10688204;
RA   Parkhill J., Wren B.W., Mungall K., Ketley J.M., Churcher C.,
RA   Basham D., Chillingworth T., Davies R.M., Feltwell T., Holroyd S.,
RA   Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA   Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA   Whitehead S., Barrell B.G.;
RT   "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT   reveals hypervariable sequences.";
RL   Nature 403:665-668(2000).
```

```
RN   [2]
RP   SEQUENCE OF 338-1031 FROM N.A.
RX   MEDLINE=96084944; PubMed=7489896;
RA   Bustamante V.H., Puente J.L., Sanchez-Lopez F., Bobadilla M.,
RA   Calva E.;
RT   "Identification of Campylobacter jejuni and C.coli using the rpoB
RT   gene and a cryptic DNA fragment from C.jejuni.";
RL   Gene 165:1-8(1995).
CC   -!- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC       OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC       SUBSTRATES.
CC   -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC       {RNA}(N).
CC   -!- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
CC       ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
CC       BETA' CHAIN.
CC   -!- SIMILARITY: Belongs to the RNA polymerase beta chain family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL139075; CAB75116.1; -.
DR   EMBL; X77304; CAA54509.1; -.
DR   PIR; A81393; A81393.
DR   PIR; S41868; S41868.
DR   HSSP; Q9KWU7; 1HQM.
DR   InterPro; IPR001572; RNA_pol_B.
DR   Pfam; PF04563; RNA_pol_Rpb2_1; 1.
DR   Pfam; PF04561; RNA_pol_Rpb2_2; 1.
DR   Pfam; PF04565; RNA_pol_Rpb2_3; 1.
DR   Pfam; PF00562; RNA_pol_Rpb2_6; 1.
DR   Pfam; PF04560; RNA_pol_Rpb2_7; 1.
DR   PROSITE; PS01166; RNA_POL_BETA; 1.
KW   Transferase; Transcription; DNA-directed RNA polymerase;
KW   Complete proteome.
FT   CONFLICT    338    347       NDLANGVDAA -> MTWLMALMQP (IN REF. 2).
FT   CONFLICT    558    558       A -> R (IN REF. 2).
FT   CONFLICT    671    671       C -> S (IN REF. 2).
FT   CONFLICT    691    691       A -> R (IN REF. 2).
SQ   SEQUENCE   1378 AA;  155915 MW;  AB7467C305028EB5 CRC64;

  Query Match             33.8%;  Score 50;  DB 1;  Length 1378;
  Best Local Similarity   40.6%;  Pred. No. 25;
  Matches   13;  Conservative    3;  Mismatches   10;  Indels    6;  Gaps    2;

Qy          3 LWAAQRYG--RELRRM----SDEFEGSFKGLK 28
              :|| : ||    || |    ||: || |   |
Db       1306 VWALEAYGAAHTLREMLTIKSDDVEGRFSAYK 1337


RESULT 13
RMUC_PSEAE
ID   RMUC_PSEAE     STANDARD;      PRT;   453 AA.
AC   Q9I4U3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA recombination protein rmuC homolog.
GN   RMUC OR PA1031.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- FUNCTION: Involved in DNA recombination (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RMUC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004535; AAG04420.1; -.
DR   PIR; E83517; E83517.
DR   InterPro; IPR003798; DUF195.
DR   Pfam; PF02646; RmuC; 1.
KW   DNA recombination; Coiled coil; Complete proteome.
FT   DOMAIN       16    201       COILED COIL (POTENTIAL).
SQ   SEQUENCE   453 AA;  51539 MW;  1E7EA97E82EC5E4B CRC64;

  Query Match             33.1%;  Score 49;  DB 1;  Length 453;
  Best Local Similarity   55.6%;  Pred. No. 11;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          4 WAAQRYGR--ELRRMSDE 19
              ||:::| ||  |||||:: |
Db         65 WASERQGREEELRRLASE 82


RESULT 14
RAS3_RHIRA
ID   RAS3_RHIRA     STANDARD;      PRT;   205 AA.
AC   P22280;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-like protein 3.
GN   RAS3.
OS   Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Mucor.
OX   NCBI_TaxID=4841;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 1216B;
RX   MEDLINE=91061774; PubMed=1701021;
RA   Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
RT   "Expression of a gene family in the dimorphic fungus Mucor racemosus
RT   which exhibits striking similarity to human ras genes.";
RL   Mol. Cell. Biol. 10:6654-6663(1990).
CC   -!- ENZYME REGULATION: ALTERNATE BETWEEN AN INACTIVE FORM BOUND TO GDP
CC       AND AN ACTIVE FORM BOUND TO GTP. ACTIVATED BY A GUANINE
CC       NUCLEOTIDE-EXCHANGE FACTOR (GEF) AND INACTIVATED BY A GTPASE-
CC       ACTIVATING PROTEIN (GAP).
CC   -!- SUBCELLULAR LOCATION: Plasma membrane.
CC   -!- DEVELOPMENTAL STAGE: IN SPORULATING MYCELIUM AND MUCH LESS IN
CC       GERMLING AND YEAST.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ---------------------------------------------------------------------
DR   EMBL; M55177; AAA83379.1; -.
DR   PIR; C36365; C36365.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      16     23       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     122    125       GTP (BY SIMILARITY).
FT   DOMAIN       38     46       EFFECTOR REGION (PROBABLE).
FT   LIPID       202    202       FARNESYL (BY SIMILARITY).
SQ   SEQUENCE   205 AA;  23408 MW;  DBF086466F090F50 CRC64;

  Query Match             32.4%;  Score 48;  DB 1;  Length 205;
  Best Local Similarity   62.5%;  Pred. No. 6.3;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         11 RELRRMSDEFEGSFKG 26
              ||:|||: | ||  ||
Db        168 REIRRMNKEQEGRSKG 183


RESULT 15
6PGL_THEMA
ID   6PGL_THEMA     STANDARD;      PRT;   220 AA.
AC   Q9X0N8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN   PGL OR DEVB OR TM1154.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC       PHOSPHOGLUCONATE.
CC   -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC       phospho-D-gluconate.
CC   -!- PATHWAY: Pentose phosphate pathway; second step.
CC   -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC       ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001772; AAD36230.1; -.
DR   PIR; F72289; F72289.
DR   TIGR; TM1154; -.
DR   InterPro; IPR006148; Gluc_gal_isom.
DR   InterPro; IPR005900; Phosphogluconlac.
DR   Pfam; PF01182; Glucosamine_iso; 1.
DR   TIGRFAMs; TIGR01198; pgl; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   220 AA;  25325 MW;  9B0FD07EE01E60C3 CRC64;

  Query Match             32.4%;  Score 48;  DB 1;  Length 220;
  Best Local Similarity   34.8%;  Pred. No. 6.8;
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AAQRYGRELRRMSDEFEGSFKGL 27
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133


Search completed: September 15, 2003, 17:23:01
Job time : 6.6 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 30.4 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-55
Perfect score: 148
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGLK 28

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	91	61.5	146	13	Q9I9N2	Q9i9n2 brachydanio
2	57	38.5	471	17	Q8ZY71	Q8zy71 pyrobaculum
3	53	35.8	196	16	Q8VJS3	Q8vjs3 mycobacteri
4	53	35.8	223	16	Q10843	Q10843 mycobacteri
5	53	35.8	946	11	Q8K016	Q8k016 mus musculu
6	52.5	35.5	505	8	O47148	O47148 menziesia c
7	52.5	35.5	506	8	O47149	O47149 rhododendro
8	52.5	35.5	506	8	O47171	O47171 rhododendro
9	52.5	35.5	506	8	O63960	O63960 rhododendro
10	52.5	35.5	506	8	O62982	O62982 rhododendro
11	52.5	35.5	506	8	O62975	O62975 rhododendro
12	52.5	35.5	506	8	O62972	O62972 rhododendro
13	52.5	35.5	506	8	O62989	O62989 rhododendro
14	52.5	35.5	506	8	O62978	O62978 rhododendro
15	52.5	35.5	506	8	O47155	O47155 rhododendro
16	52.5	35.5	506	8	O47152	O47152 rhododendro
17	52.5	35.5	506	8	O47173	O47173 rhododendro
18	52.5	35.5	506	8	O62990	O62990 rhododendro
19	52.5	35.5	506	8	O62974	O62974 rhododendro
20	52.5	35.5	506	8	O62993	O62993 menziesia m
21	52.5	35.5	506	8	O47170	O47170 rhododendro
22	52.5	35.5	506	8	O47174	O47174 rhododendro
23	52.5	35.5	506	8	O62983	O62983 rhododendro
24	52.5	35.5	506	8	O62980	O62980 rhododendro
25	52.5	35.5	506	8	O62981	O62981 rhododendro
26	52.5	35.5	506	8	O62977	O62977 rhododendro
27	52.5	35.5	506	8	O47168	O47168 menziesia p
28	52.5	35.5	506	8	O62988	O62988 rhododendro
29	52.5	35.5	506	8	O62973	O62973 rhododendro
30	52.5	35.5	506	8	O62992	O62992 ledum palus
31	52.5	35.5	506	8	O47175	O47175 rhododendro
32	52.5	35.5	506	8	Q8HSP1	Q8hspl rhododendro
33	52.5	35.5	506	8	Q8HSP0	Q8hsp0 rhododendro
34	52.5	35.5	506	8	Q8HSN9	Q8hsn9 rhododendro
35	52.5	35.5	506	8	Q8HSN8	Q8hsn8 rhododendro
36	52.5	35.5	506	8	Q8HSN7	Q8hsn7 rhododendro
37	52.5	35.5	506	8	Q8HSN6	Q8hsn6 rhododendro
38	52.5	35.5	506	8	Q8HSN5	Q8hsn5 rhododendro
39	52.5	35.5	506	8	Q8HSN4	Q8hsn4 rhododendro
40	52.5	35.5	507	8	O62985	O62985 rhododendro
41	52.5	35.5	507	8	O62986	O62986 rhododendro
42	52.5	35.5	508	8	O62979	O62979 rhododendro
43	51.5	34.8	506	8	O47153	O47153 rhododendro
44	51.5	34.8	506	8	O47160	O47160 rhododendro
45	51.5	34.8	506	8	Q8HSP4	Q8hsp4 rhododendro

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 61.5%; Score 91; DB 13; Length 146;
Best Local Similarity 61.5%; Pred. No. 2.1e-05;
Matches 16; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

```
Qy          3 LWAAQRYGRELRRMSDEFEGSFKGLK 28
              ||||::||::||||||||:   | :|
Db         89 LWAAKKYGQQLRRMSDEFDKGMKRVK 114
```

RESULT 2
Q8ZY71
ID Q8ZY71 PRELIMINARY; PRT; 471 AA.
AC Q8ZY71;

```
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein PAE0922.
GN   PAE0922.
OS   Pyrobaculum aerophilum.
OC   Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC   Thermoproteaceae; Pyrobaculum.
OX   NCBI_TaxID=13773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX   PubMed=11792869;
RA   Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA   Miller J.H.;
RT   "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT   aerophilum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR   EMBL; AE009793; AAL63125.1; -.
DR   InterPro; IPR006638; Elp3.
DR   InterPro; IPR000182; GCN5acetyltransf.
DR   Pfam; PF00583; Acetyltransf; 1.
DR   SMART; SM00729; Elp3; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   471 AA;  52952 MW;  3B1E36E8AEE2EF0A CRC64;

  Query Match             38.5%;  Score 57;  DB 17;  Length 471;
  Best Local Similarity   44.0%;  Pred. No. 7;
  Matches   11;  Conservative    3;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          4 WAAQRYGRELRRMSDEFEGSFKGLK 28
              |     |||| |:::|  | |  ||
Db        404 WQHSGMGRELMRLAEEIAGEFGALK 428


RESULT 3
Q8VJS3
ID   Q8VJS3      PRELIMINARY;      PRT;   196 AA.
AC   Q8VJS3;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   IS1607, transposase.
GN   MT2070.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE007058; AAK46348.1; -.
DR   TIGR; MT2070; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
SQ   SEQUENCE   196 AA;  21349 MW;  C145A8D836FD9C2D CRC64;

  Query Match             35.8%;  Score 53;  DB 16;  Length 196;
  Best Local Similarity   58.8%;  Pred. No. 10;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              |||| || | : |  |
Db        134 NLWAADRYNRAIARGHD 150


RESULT 4
Q10843
ID   Q10843      PRELIMINARY;      PRT;   223 AA.
AC   Q10843;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein Rv2014.
GN   RV2014 OR MTCY39.03C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SIMILARITY: TO M.PARATUBERCULOSIS IS900.
DR   EMBL; Z74025; CAA98415.1; -.
DR   TubercuList; Rv2014; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   223 AA;  24132 MW;  70456750017FEF37 CRC64;

  Query Match             35.8%;  Score 53;  DB 16;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 12;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              |||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 5
Q8K016
ID   Q8K016      PRELIMINARY;      PRT;   946 AA.
AC   Q8K016;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Inter-alpha trypsin inhibitor, heavy chain 2.
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC034341; AAH34341.1; -.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE   946 AA;  105945 MW;  8B17DBA71B85BC5C CRC64;
```

```
  Query Match           35.8%;  Score 53;  DB 11;  Length 946;
  Best Local Similarity  34.6%;  Pred. No. 58;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |    : | ||| |:|:
Db        212 NVWIMEPQGMRFLHVPDTFEGHFQGV 237


RESULT 6
O47148
ID   O47148      PRELIMINARY;      PRT;   505 AA.
AC   O47148;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K) (Fragment).
GN   MATK.
OS   Menziesia ciliicalyx.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Menziesia.
OX   NCBI_TaxID=49154;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61331; AAC15245.2; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   505 AA;  60233 MW;  EE5F927AD2E57DE5 CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 505;
  Best Local Similarity  37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        390 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 421


RESULT 7
O47149
ID   O47149      PRELIMINARY;      PRT;   506 AA.
AC   O47149;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron kiusianum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49167;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61332; AAB93753.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60502 MW;  0009EA88CD28549F CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity  37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 8
O47171
ID   O47171      PRELIMINARY;      PRT;   506 AA.
AC   O47171;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron edgeworthii.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49162;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61354; AAB93748.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60485 MW;  8A6353BFC5F4DC85 CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity  37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 9
O63960
ID   O63960      PRELIMINARY;      PRT;   506 AA.
AC   O63960;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK OR YCF14.
OS   Rhododendron tashiroi, and
```

```
OS   Rhododendron farrerae.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75582, 75580;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012749; BAA25870.1; -.
DR   EMBL; AB012745; BAA25866.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60389 MW;  DE0C07AEE608B787 CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 10
O62982
ID   O62982      PRELIMINARY;      PRT;   506 AA.
AC   O62982;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron nipponicum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012739; BAA25860.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60419 MW;  1F95132CCF4F6B40 CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 11
O62975
ID   O62975      PRELIMINARY;      PRT;   506 AA.
AC   O62975;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron ponticum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49628;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012732; BAA25853.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60449 MW;  21DFF700B071B5B8 CRC64;

  Query Match           35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 12
O62972
ID   O62972      PRELIMINARY;      PRT;   506 AA.
AC   O62972;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron ovatum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49169;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
```

```
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012729; BAA25850.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60493 MW;  D230E54B8C20FEF0 CRC64;

  Query Match             35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 13
O62989
ID   O62989      PRELIMINARY;      PRT;   506 AA.
AC   O62989;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron indicum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75581;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012747; BAA25868.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60489 MW;  6D38A1D4D6FEC9BF CRC64;

  Query Match             35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 14
O62978
ID   O62978      PRELIMINARY;      PRT;   506 AA.
AC   O62978;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron canadense.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49465;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012735; BAA25856.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60350 MW;  5E832589ED64EA25 CRC64;

  Query Match             35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 15
O47155
ID   O47155      PRELIMINARY;      PRT;   506 AA.
AC   O47155;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron hongkongense.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49165;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61338; AAB93751.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60406 MW;  4B5C675CE32218D8 CRC64;

  Query Match             35.5%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 34;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
```

```
                    | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422

Search completed: September 15, 2003, 17:25:50
Job time : 31.4 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 38.1857 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	143	100.0	27	21	AAB37056	Bcl2 polypeptide B
2	143	100.0	28	21	AAB37055	Bcl2 polypeptide B
3	138	96.5	26	21	AAB37001	Bcl2 polypeptide B
4	138	96.5	26	21	AAB37002	Bcl2 polypeptide B
5	138	96.5	27	21	AAB37003	Bcl2 polypeptide B
6	138	96.5	162	22	AAB70370	Shorter murine BAD
7	138	96.5	204	17	AAR95168	bcl-x(L)/bcl-2 ass
8	138	96.5	204	19	AAW61315	Murine BCL-XL/BCL-
9	138	96.5	204	19	AAW61316	Mutant BCL-XL/BCL-
10	138	96.5	204	19	AAW61317	Mutant BCL-XL/BCL-
11	138	96.5	204	19	AAW61318	Mutant BCL-XL/BCL-
12	138	96.5	204	19	AAW58832	Murine BAD protein
13	138	96.5	204	22	AAB70369	Longer murine BAD
14	138	96.5	204	24	ABR39082	Murine BAD protein
15	138	96.5	567	22	AAU00220	Bad-DTTR apoptosis
16	114	79.7	24	23	AAU78627	Human Bad peptide
17	114	79.7	25	23	ABP56161	PTPC-interacting T
18	114	79.7	25	23	ABG78484	Mutant Bcl2 compet
19	114	79.7	25	23	ABG78493	Mutant Bcl2 compet
20	114	79.7	25	23	AAU78610	Human Bad peptide
21	114	79.7	25	23	AAU78620	Human Bad peptide
22	114	79.7	166	18	AAW32476	BBC6 protein for r
23	114	79.7	168	19	AAW55779	Human Bcl-xL/Bcl-2
24	114	79.7	168	21	AAB13512	Human cell prolife
25	114	79.7	168	22	AAB70368	Human BAD mutant a
26	114	79.7	168	22	AAB48287	Human Bad protein.
27	114	79.7	168	22	AAG67688	Amino acid sequenc
28	114	79.7	168	24	ABR39081	Human BAD protein
29	114	79.7	201	23	ABP41630	Human ovarian anti
30	113	79.0	23	17	AAR95166	bcl-x(L)/bcl-2 ass
31	111	77.6	25	23	ABG78490	Mutant Bcl2 compet
32	111	77.6	25	23	AAU78617	Human Bad peptide
33	110	76.9	25	23	ABG78488	Mutant Bcl2 compet
34	110	76.9	25	23	ABG78489	Mutant Bcl2 compet
35	110	76.9	25	23	AAU78615	Human Bad peptide
36	110	76.9	25	23	AAU78616	Human Bad peptide
37	109	76.2	23	23	AAU78628	Human Bad peptide
38	109	76.2	25	23	ABG78486	Mutant Bcl2 compet
39	109	76.2	25	23	ABG78492	Mutant Bcl2 compet
40	109	76.2	25	23	ABG78497	Mutant Bcl2 compet
41	109	76.2	25	23	AAU78612	Human Bad peptide
42	109	76.2	25	23	AAU78619	Human Bad peptide
43	109	76.2	25	23	AAU78624	Human Bad peptide
44	108	75.5	25	23	ABG78485	Mutant Bcl2 compet
45	108	75.5	25	23	AAU78611	Human Bad peptide

ALIGNMENTS

RESULT 1
AAB37056
ID AAB37056 standard; peptide; 27 AA.
XX
AC AAB37056;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #56.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.

XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 143; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 7.3e-15;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||||
Db 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 2
AAB37055
ID AAB37055 standard; peptide; 28 AA.
XX
AC AAB37055;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #55.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 28 AA;

Query Match 100.0%; Score 143; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 7.6e-15;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||||
Db 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

RESULT 3
AAB37001
ID AAB37001 standard; peptide; 26 AA.
XX
AC AAB37001;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #1.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.

XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 96.5%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No.-4.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 4
AAB37002
ID AAB37002 standard; peptide; 26 AA.
XX
AC AAB37002;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #2.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 96.5%; Score 138; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.1e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 5
AAB37003
ID AAB37003 standard; peptide; 27 AA.
XX
AC AAB37003;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #3.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.

XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 96.5%; Score 138; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.3e-14;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 1 NLWAAQRYGRELRRMSDEFEGSFKGL 26

RESULT 6
AAB70370
ID AAB70370 standard; protein; 162 AA.
XX
AC AAB70370;
XX
DT 02-MAY-2001 (first entry)
XX
DE Shorter murine BAD mutant amino acid sequence SEQ ID NO:3.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 7; Page 148-149; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed shorter murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 162 AA;

Query Match 96.5%; Score 138; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.1e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 98 NLWAAQRYGRELRRMSDEFEGSFKGL 123

RESULT 7
AAR95168
ID AAR95168 standard; Protein; 204 AA.
XX
AC AAR95168;
XX
DT 06-JAN-1997 (first entry)
XX
DE bcl-x(L)/bcl-2 associated death promoter protein.
XX
KW Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS Mus musculus.
XX

```
FH   Key             Location/Qualifiers
FT   Region          147..149
FT                   /note= "BH1 conserved amino acids"
FT   Region          191..192
FT                   /note= "BH2 conserved amino acids"
FT   Domain          38..61
FT                   /note= "PEST sequence"
FT   Domain          111..130
FT                   /note= "PEST sequence"
XX
PN   WO9613614-A1.
XX
PD   09-MAY-1996.
XX
PF   31-OCT-1995;   95WO-US14246.
XX
PR   31-OCT-1994;   94US-0333565.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1996-251465/25.
DR   N-PSDB; AAT29479.
XX
PT   Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT   useful to treat neoplasia and apoptosis and to identify agents
PT   inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS   Claim 3; Fig 1; 130pp; English.
XX
CC   This sequence represents the murine bcl-x(L)/bcl-2 associated death
CC   promoter (Bad) gene.  Bad is a 22.1 kD protein which interacts with
CC   bcl-2 and bcl-x proteins and regulates cell death.  It has homology
CC   to the bcl-2-related family clustered in the BH1 and BH2 domain.  Bad
CC   has been found to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid
CC   assays and in vivo in mammalian cells.  Overexpressed Bad counters the
CC   death inhibitory activity of bcl-x(L), but is much less effective at
CC   countering the death inhibitory activity of bcl-2.  Bad expression can
CC   accelerate apoptotic cell death induced by cytokine deprivation in an
CC   IL-3 dependent cell line expressing bcl-x(L), and its also counters the
CC   death repressor activity of bcl-x(L).  Bad competes with Bax for binding
CC   to bcl-x(L).  Bad may be used to identify agents which inhibit its
CC   binding to bcl-2 or bcl-x(L) to form heterodimers.  Such agents may be
CC   used to treat neurodegenerative diseases, immunodeficiency diseases,
CC   e.g. AIDS, senescence or ischaemia.
XX
SQ   Sequence   204 AA;

  Query Match             96.5%;  Score 138;  DB 17;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 8
AAW61315
ID   AAW61315 standard; Protein; 204 AA.
XX
AC   AAW61315;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Murine BCL-XL/BCL-2 associated cell death regulator.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27833.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 1; Fig 10; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence is the murine BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             96.5%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 9
AAW61316
ID   AAW61316 standard; Protein; 204 AA.
XX
AC   AAW61316;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #1.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
```

XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27834.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 59; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 96.5%; Score 138; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 10
AAW61317
ID AAW61317 standard; Protein; 204 AA.
XX
AC AAW61317;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #2.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27835.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 60; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 96.5%; Score 138; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 11
AAW61318
ID AAW61318 standard; Protein; 204 AA.
XX
AC AAW61318;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #3.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.

```
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27836.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 60-61; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             96.5%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 12
AAW58832
ID   AAW58832 standard; protein; 204 AA.
XX
AC   AAW58832;
XX
DT   23-JUL-1998  (first entry)
XX
DE   Murine BAD protein.
XX
KW   BAD protein; Bcl-XL/Bcl-2 associated cell death regulator; 14-3-3;
KW   serine phosphorylation; post-translational modification; apoptosis;
KW   signal transduction regulator; phosphoserine phosphatase; senescence;
KW   immunodeficiency disease; neurodegenerative disease; infertility;
KW   cancer, viral infection; lymphoproliferative condition; arthritis;
KW   inflammation; autoimmune diseases.
XX
OS   Mus sp.
XX
PN   WO9809643-A1.
XX
PD   12-MAR-1998.
XX
PF   09-SEP-1997;   97WO-US15871.
XX
PR   09-SEP-1996;   96US-0707868.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-207049/18.
XX
PT   Serine-phosphorylated Bcl-X-l/Bcl-2 Associated cell Death regulator
PT   polypeptide - useful for modulation of apoptosis associated with,
PT   e.g. cancer and immunodeficiency diseases
XX
PS   Claim 3; Fig 8; 61pp; English.
XX
CC   This sequence represents a novel serine-phosphorylated protein, BAD
CC   (Bcl-Xl/Bcl-2 associated cell death regulator). The serine residue is
CC   phosporylated in a post-translational modification and allows binding
CC   to the 14-3-3 protein which is a signal transduction regulator.
CC   Modulators of phosphorylated BAD, which act through inhibition/activation
CC   of a phosphoserine phosphatase, are useful for preventing/treating
CC   increased/decreased apoptosis in a cell. The increased apoptosis may
CC   result from immunodeficiency diseases, senescence, neurodegenerative
CC   disease, ischaemic cell death, reperfusion cell death, infertility and
CC   wound-healing.  Decreased apoptosis may result from cancer, viral
CC   infection, lymphoproliferative conditions, arthritis, infertility,
CC   inflammation and autoimmune diseases. Measuring the amount of
CC   phosphorylated compared to unphosphorylated BAD polypeptide and/or total
CC   BAD in a cell is useful for determining the apoptotic state of a cell.
XX
SQ   Sequence   204 AA;

  Query Match             96.5%;  Score 138;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 4e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 13
AAB70369
ID   AAB70369 standard; protein; 204 AA.
XX
AC   AAB70369;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Longer murine BAD mutant amino acid sequence SEQ ID NO:2.
XX
KW   Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW   immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW   cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW   immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW   immunodeficiency disease; neurodegenerative disease; viral infection;
KW   ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW   lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS   Mus musculus.
OS   Synthetic.
XX
PN   WO200110888-A1.
XX
PD   15-FEB-2001.
XX
PF   30-MAY-2000; 2000WO-US11864.
XX
PR   28-MAY-1999;   99US-0136783.
```

XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit
PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 4; Page 148; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed longer murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 204 AA;

Query Match 96.5%; Score 138; DB 22; Length 204;
Best Local Similarity 100.0%; Pred. No. 4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 14
ABR39082
ID ABR39082 standard; Protein; 204 AA.
XX
AC ABR39082;
XX
DT 10-MAY-2003 (first entry)
XX
DE Murine BAD protein SEQ ID NO:4.
XX
KW Murine; BAD; herpes simplex virus; HSV; US3; herpes virus; apoptosis;
KW virucide; infection.
XX
OS Mus musculus.
XX
PN WO2003012049-A2.
XX
PD 13-FEB-2003.
XX
PF 31-JUL-2002; 2002WO-US24177.
XX
PR 31-JUL-2001; 2001US-308929P.
XX
PA (UYCH-) UNIV CHICAGO.
XX
PI Munger J, Roizman B;
XX
DR WPI; 2003-248168/24.
DR N-PSDB; ABZ81201.
XX
PT Inducing apoptosis in a cell infected with herpes simplex virus, HSV,
PT by administering to the cell, a composition comprising an agent that
PT inhibits phosphorylation of pro-apoptotic polypeptide BAD by HSV US3 -
XX
PS Claim 15; Page 168; 192pp; English.
XX
CC The present invention describes a method (M1) for inducing apoptosis in
CC a cell infected with herpes simplex virus (HSV), which comprises
CC administering to the cell, a composition having an agent that inhibits
CC phosphorylation of pro-apoptotic polypeptide BAD by HSV US3. Also
CC described is a method (M2) for treating a patient infected with HSV, by
CC administering to the patient, a composition comprising a peptide
CC comprising a sequence of 4-100 continuous amino acids of a 168 residue
CC amino acid sequence (see ABR39081), where the peptide comprises ser112,
CC ser135, or ser155, or their combinations. BAD has virucide activity.
CC M1 is useful for inducing apoptosis in a cell infected with HSV, where
CC the cell is in a human. M2 is useful for treating a patient infected
CC with HSV. The present sequence represents murine BAD, which is used in
CC the exemplification of the present invention.
XX
SQ Sequence 204 AA;

Query Match 96.5%; Score 138; DB 24; Length 204;
Best Local Similarity 100.0%; Pred. No. 4e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 15
AAU00220
ID AAU00220 standard; Protein; 567 AA.
XX
AC AAU00220;
XX
DT 31-MAY-2001 (first entry)
XX
DE Bad-DTTR apoptosis-modifying fusion protein.
XX
KW Mouse; Bad-DTTR; apoptosis; cancer; spinal muscular atrophy;
KW diphtheria toxin receptor binding domain; DTR; neoplasm; tumour;
KW hyper-proliferation; Alzheimer's disease; neurodegenerative disorder;
KW transient ischaemic neuronal injury; stroke; spinal cord injury;
KW Huntington's disease.
XX
OS Chimeric - Mus sp.
OS Chimeric - Corynebacterium diptheriae.
OS Chimeric - Synthetic.
XX
FH Key Location/Qualifiers
FT Region 3..12
FT /note= "10x histidine tag"
XX
PN WO200112661-A2.
XX
PD 22-FEB-2001.
XX
PF 15-AUG-2000; 2000WO-US22293.
XX
PR 16-AUG-1999; 99US-0149220.
XX
PA (HARD) HARVARD COLLEGE.
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Youle RJ, Liu X, Collier RJ;
XX
DR WPI; 2001-218343/22.
DR N-PSDB; AAS00248.
XX

PT Novel fusion protein for modifying apoptosis in target cell and
PT reducing apoptosis after transient ischaemic neuronal injury, has two
PT domains which targets protein to a cell and modifies apoptotic response
PT of cell -
XX
PS Claim 4; Page 59-61; 65pp; English.
XX
CC The sequence represents the amino acid sequence of Bad-DTTR apoptosis-
CC modifying fusion protein comprising Bad gene sequence fused via a short
CC linker to diphtheria toxin translocation domain (DTTR). The
CC functional apoptosis-modifying fusion protein is capable of binding a
CC target cell and integrating into or crossing a cellular membrane of the
CC target cell.. The apoptosis-modifying fusion protein comprises at least
CC two domains: the DTR domain, which targets the fusion protein to the
CC target cell and the Bcl-XL domain, which modifies an apoptotic response
CC of the target cell. The fusion protein is useful for modifying
CC (inhibiting or enhancing) apoptosis in a target cell, such as neuron,
CC lymphocyte, cancer, neoplasm, macrophage, epithelial, stem, tumour or
CC hyper-proliferative cell or an adipocyte. It is also useful for reducing
CC apoptosis in a subject after transient ischaemic neuronal injury,
CC especially spinal cord injury. The fusion protein may be used to treat
CC various diseases and injury conditions through inhibition or enhancement
CC of apoptotic cellular response, including neurodegenerative disorders
CC such as Alzheimer's disease, Huntington's disease, spinal muscular
CC atrophy, stroke episodes and unregulated cell growth as in tumours and
CC various cancers. The apoptosis-modifying fusion protein can be delivered
CC effectively throughout the body and targeted to selective tissue and
CC cells.
XX
SQ Sequence 567 AA;

Query Match 96.5%; Score 138; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 1.2e-12;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        161 NLWAAQRYGRELRRMSDEFEGSFKGL 186
```

Search completed: September 15, 2003, 17:22:14
Job time : 38.1857 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 14.0786 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	96.5	204	1	US-08-333-565-2	Sequence 2, Appli
2	138	96.5	204	2	US-08-661-479-2	Sequence 2, Appli
3	138	96.5	204	2	US-08-733-505A-1	Sequence 1, Appli
4	138	96.5	204	2	US-08-733-505A-12	Sequence 12, Appl
5	138	96.5	204	2	US-08-733-505A-13	Sequence 13, Appl
6	138	96.5	204	2	US-08-733-505A-14	Sequence 14, Appl
7	135	94.4	204	2	US-08-717-123-3	Sequence 3, Appli
8	135	94.4	204	4	US-09-375-257-3	Sequence 3, Appli
9	114	79.7	166	1	US-08-665-617-2	Sequence 2, Appli
10	114	79.7	168	2	US-08-717-123-2	Sequence 2, Appli
11	114	79.7	168	3	US-08-985-335-1	Sequence 1, Appli
12	114	79.7	168	3	US-08-985-335-7	Sequence 7, Appli
13	114	79.7	168	3	US-09-410-372-1	Sequence 1, Appli
14	114	79.7	168	3	US-09-410-372-7	Sequence 7, Appli
15	114	79.7	168	4	US-09-375-257-2	Sequence 2, Appli
16	113	79.0	23	1	US-08-333-565-10	Sequence 10, Appl
17	113	79.0	23	2	US-08-661-479-10	Sequence 10, Appl
18	102	71.3	59	2	US-08-733-505A-55	Sequence 55, Appl
19	102	71.3	59	2	US-08-733-505A-56	Sequence 56, Appl
20	102	71.3	59	2	US-08-733-505A-57	Sequence 57, Appl
21	102	71.3	59	2	US-08-733-505A-58	Sequence 58, Appl
22	86	60.1	16	1	US-08-333-565-26	Sequence 26, Appl
23	86	60.1	16	2	US-08-661-479-26	Sequence 26, Appl
24	61	42.7	11	2	US-08-733-505A-34	Sequence 34, Appl
25	61	42.7	11	2	US-08-706-741B-69	Sequence 69, Appl
26	61	42.7	11	2	US-08-924-695A-69	Sequence 69, Appl
27	51	35.7	66	2	US-08-867-087B-40	Sequence 40, Appl
28	48.5	33.9	904	4	US-09-328-352-4656	Sequence 4656, Ap
29	46	32.2	610	4	US-09-252-991A-19594	Sequence 19594, A
30	46	32.2	946	3	US-09-074-579-3	Sequence 3, Appli
31	46	32.2	946	3	US-09-388-774-3	Sequence 3, Appli
32	46	32.2	946	4	US-09-546-153-1	Sequence 1, Appli
33	45.5	31.8	906	4	US-09-252-991A-31458	Sequence 31458, A
34	45	31.5	229	4	US-09-252-991A-23807	Sequence 23807, A
35	45	31.5	303	4	US-09-328-352-5164	Sequence 5164, Ap
36	45	31.5	356	4	US-09-235-103-2	Sequence 2, Appli
37	45	31.5	356	4	US-09-235-103-4	Sequence 4, Appli
38	45	31.5	1064	4	US-09-252-991A-17508	Sequence 17508, A
39	44.5	31.1	903	4	US-09-252-991A-28775	Sequence 28775, A
40	44	30.8	125	4	US-09-328-352-7449	Sequence 7449, Ap
41	44	30.8	263	4	US-09-651-656-27	Sequence 27, Appl
42	44	30.8	263	4	US-09-650-855-27	Sequence 27, Appl
43	44	30.8	277	4	US-09-252-991A-28581	Sequence 28581, A
44	44	30.8	877	4	US-09-206-551-20	Sequence 20, Appl
45	44	30.8	1125	4	US-09-252-991A-18729	Sequence 18729, A

ALIGNMENTS

```
RESULT 1
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-333-565-2

  Query Match           96.5%;  Score 138;  DB 1;  Length 204;
  Best Local Similarity  100.0%;  Pred. No. 1e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-717-123-3

  Query Match             94.4%;  Score 135;  DB 2;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 2.9e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 8
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match             94.4%;  Score 135;  DB 4;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 2.9e-13;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 9
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match             79.7%;  Score 114;  DB 1;  Length 166;
  Best Local Similarity   91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||| |||
Db        101 NLWAAQRYGRELRRMSDEFVDSFK 124


RESULT 10
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match             79.7%;  Score 114;  DB 2;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||| |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126
```

```
RESULT 11
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match            79.7%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 12
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match            79.7%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 13
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
```

```
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match           79.7%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 14
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match           79.7%;  Score 114;  DB 3;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 15
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match           79.7%;  Score 114;  DB 4;  Length 168;
  Best Local Similarity  91.7%;  Pred. No. 4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


Search completed: September 15, 2003, 17:45:07
Job time : 15.0786 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 21.4071 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	135	94.4	204	9	US-09-922-378-3	Sequence 3, Appli
2	135	94.4	204	14	US-10-066-179-3	Sequence 3, Appli
3	114	79.7	25	15	US-10-059-261-258	Sequence 258, App
4	114	79.7	168	9	US-09-922-378-2	Sequence 2, Appli
5	114	79.7	168	9	US-09-894-657-1	Sequence 1, Appli
6	114	79.7	168	9	US-09-894-657-7	Sequence 7, Appli
7	114	79.7	168	14	US-10-066-179-2	Sequence 2, Appli
8	71	49.7	15	15	US-10-174-105A-147	Sequence 147, App
9	51	35.7	682	12	US-10-238-075-1077	Sequence 1077, Ap
10	50	35.0	215	15	US-10-156-761-9145	Sequence 9145, Ap
11	47	32.9	35	15	US-10-092-750-1	Sequence 1, Appli
12	47	32.9	138	15	US-10-092-750-241	Sequence 241, App
13	46	32.2	946	9	US-09-828-423-3	Sequence 3, Appli
14	44	30.8	272	15	US-10-156-761-11541	Sequence 11541, A
15	44	30.8	426	9	US-09-815-242-5704	Sequence 5704, Ap
16	44	30.8	699	14	US-10-008-355-8	Sequence 8, Appli
17	44	30.8	705	9	US-09-815-242-12463	Sequence 12463, A
18	44	30.8	712	14	US-10-008-355-2	Sequence 2, Appli
19	44	30.8	877	12	US-10-369-294-20	Sequence 20, Appl
20	43	30.1	213	9	US-09-843-846-18	Sequence 18, Appl
21	43	30.1	232	10	US-09-881-752A-238	Sequence 238, App
22	43	30.1	270	11	US-09-934-455-162	Sequence 162, App
23	43	30.1	380	9	US-09-149-045-2	Sequence 2, Appli
24	43	30.1	380	15	US-10-166-359-2	Sequence 2, Appli
25	43	30.1	380	15	US-10-166-113-2	Sequence 2, Appli
26	43	30.1	380	15	US-10-166-357-2	Sequence 2, Appli
27	43	30.1	380	15	US-10-166-372-2	Sequence 2, Appli
28	43	30.1	380	15	US-10-184-722-3	Sequence 3, Appli
29	43	30.1	380	15	US-10-251-385-62	Sequence 62, Appl
30	43	30.1	380	15	US-10-251-385-198	Sequence 198, App
31	43	30.1	380	15	US-10-225-567A-233	Sequence 233, App
32	43	30.1	543	15	US-10-156-761-13485	Sequence 13485, A
33	43	30.1	571	9	US-09-815-242-11813	Sequence 11813, A
34	43	30.1	582	10	US-09-331-631A-22	Sequence 22, Appl
35	43	30.1	640	9	US-09-989-722-501	Sequence 501, App
36	43	30.1	640	9	US-09-989-723-501	Sequence 501, App
37	43	30.1	640	9	US-09-989-279-501	Sequence 501, App
38	43	30.1	640	9	US-09-989-727-501	Sequence 501, App
39	43	30.1	640	10	US-09-989-731-501	Sequence 501, App
40	43	30.1	640	10	US-09-989-732-501	Sequence 501, App
41	43	30.1	640	10	US-09-991-073-501	Sequence 501, App
42	43	30.1	640	10	US-09-909-320-292	Sequence 292, App
43	43	30.1	640	10	US-09-990-442-501	Sequence 501, App
44	43	30.1	640	10	US-09-991-163-501	Sequence 501, App
45	43	30.1	640	10	US-09-993-604-501	Sequence 501, App

ALIGNMENTS

RESULT 1
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE
; FILE REFERENCE: 480140.428D3
; CURRENT APPLICATION NUMBER: US/09/922,378
; CURRENT FILING DATE: 2001-08-03
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 204
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-922-378-3

Query Match 94.4%; Score 135; DB 9; Length 204;
Best Local Similarity 96.2%; Pred. No. 5.5e-12;
Matches 25; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||:|||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165
```

RESULT 2
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
; APPLICANT: Horne, William A.
; APPLICANT: Oltersdorf, Tilman
; TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
; TITLE OF INVENTION: ACIDS AND METHODS OF USE

```
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             94.4%;  Score 135;  DB 14;  Length 204;
  Best Local Similarity   96.2%;  Pred. No. 5.5e-12;
  Matches   25;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |||||||||||||||:||||||||||
Db        140 NLWAAQRYGRELRRMTDEFEGSFKGL 165


RESULT 3
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match             79.7%;  Score 114;  DB 15;  Length 25;
  Best Local Similarity   91.7%;  Pred. No. 7.4e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||||  |||
Db          1 NLWAAQRYGRELRRMSDEFVDSFK 24


RESULT 4
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match             79.7%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.1e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 5
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;  INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: SYNORAB01
;            CLONE: 358673
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             79.7%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.1e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              ||||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 6
US-09-894-657-7
```

```
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: GenBank
;            CLONE: 1683637
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match             79.7%;  Score 114;  DB 9;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.1e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||| |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 7
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match             79.7%;  Score 114;  DB 14;  Length 168;
  Best Local Similarity   91.7%;  Pred. No. 5.1e-09;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFK 25
              |||||||||||||||||||| |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPEC
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCH
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match             49.7%;  Score 71;  DB 15;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.0008;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy         10 GRELRRMSDEFEGS 23
              ||||||||||||||
Db          1 GRELRRMSDEFEGS 14


RESULT 9
US-10-238-075-1077
; Sequence 1077, Application US/10238075
; Publication No. US20030148324A1
; GENERAL INFORMATION:
;  APPLICANT: I.N.S.E.R.M.
;  TITLE OF INVENTION: Polynucleotides which are of nature B2/D+ A- and which are iso
;  TITLE OF INVENTION:  E.coli, and biological uses of these polynucleotides and of t
;  FILE REFERENCE: BLANDINE
;  CURRENT APPLICATION NUMBER: US/10/238,075
;  CURRENT FILING DATE:  2002-09-10
;  PRIOR APPLICATION NUMBER: 0003145
;  PRIOR FILING DATE: 2000-03-10
;  NUMBER OF SEQ ID NOS: 1576
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1077
;   LENGTH: 682
;   TYPE: PRT
;   ORGANISM: Escherichia coli
```

US-10-238-075-1077

Query Match 35.7%; Score 51; DB 12; Length 682;
Best Local Similarity 40.7%; Pred. No. 32;
Matches 11; Conservative 5; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27
              | :||||| | :: |: : |     ||
Db        608 KQIWAAQRNGAKVPRVRNGFTSMDIGL 634
```

RESULT 10
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

Query Match 35.0%; Score 50; DB 15; Length 215;
Best Local Similarity 56.2%; Pred. No. 14;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```
Qy          7 QRYGRELRRMSDEFEG 22
              :|:| :|||| || :|
Db        108 ERWGGDLRRMRDEADG 123
```

RESULT 11
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 35
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-1

Query Match 32.9%; Score 47; DB 15; Length 35;
Best Local Similarity 45.5%; Pred. No. 5.9;
Matches 10; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

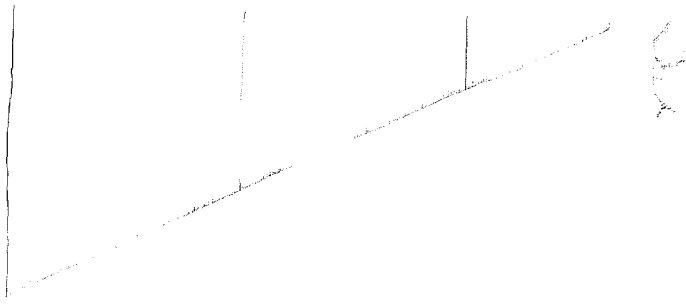
```
Qy          3 LWAAQRYGRELRRMSDEFEGSF 24
              :| ||    ||||: |||   :
Db         15 IWIAQ----ELRRIGDEFNAYY 32
```

RESULT 12
US-10-092-750-241
; Sequence 241, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 241
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-241

Query Match 32.9%; Score 47; DB 15; Length 138;
Best Local Similarity 45.5%; Pred. No. 24;
Matches 10; Conservative 3; Mismatches 5; Indels 4; Gaps 1;

```
Qy          3 LWAAQRYGRELRRMSDEFEGSF 24
              :| ||    ||||: |||   :
Db         86 IWIAQ----ELRRIGDEFNAYY 103
```

RESULT 13
US-09-828-423-3
; Sequence 3, Application US/09828423
; Patent No. US20020099178A1
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; Guegler, Karl J.
; Patterson, Chandra
; TITLE OF INVENTION: GROWTH-ASSOCIATED TRYPSIN-TYPE
; INHIBITOR HEAVY CHAIN PRECURSOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: Word Perfect 6.1/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/828,423
; FILING DATE: 05-Apr-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/388,774
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Cerrone, Michael C
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0505 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166



; 3 ;

core 46; DB 9; Length 946;
Pred. No. 2.4e+02;
5; Mismatches 13; Indels 0; Gaps 0;

Qy SDEFEGSFKGL 27
: | ||| | |:
Db LHVPDTFEGHFDGV 237

RESULT 14
US-10-156-761-11S
; Sequence 11541, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 11541
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-11541

Query Match 30.8%; Score 44; DB 15; Length 272;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 4 WAAQRYGRELRRMSD 18
| | :| |||| :|
Db 29 WIAAAHGAELRRAAD 43

RESULT 15
US-09-815-242-5704
; Sequence 5704, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5704
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5704

Query Match 30.8%; Score 44; DB 9; Length 426;
Best Local Similarity 36.8%; Pred. No. 2.1e+02;
Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 9 YGRELRRMSDEFEGSFKGL 27
:| ||| : |:| | :
Db 386 FGGSLRRQDENFDGKIKAI 404

Search completed: September 15, 2003, 17:47:53
Job time : 21.4071 secs

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 12.15 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	96.5	204	2	A55671	bad protein - mous
2	54	37.8	946	2	JC5575	inter-alpha-trypsi
3	53	37.1	223	2	D70760	hypothetical prote
4	53	37.1	946	2	S54354	inter-alpha-inhibi
5	52	36.4	370	2	S38185	2-dehydro-3-deoxy-
6	51	35.7	232	2	A42095	floral homeotic pr
7	50	35.0	374	2	C84338	spermidine/putresc
8	50	35.0	516	2	A96753	probable threonine
9	49	34.3	453	2	E83517	conserved hypothet
10	48.5	33.9	134	2	S40376	Ig kappa chain - h
11	48.5	33.9	314	2	T02975	annexin P35 - maiz
12	48	33.6	206	2	C36365	transforming prote
13	48	33.6	220	2	F72289	oxidoreductase, so
14	48	33.6	526	2	T08545	threonine synthase
15	47.5	33.2	334	2	A39172	Antho-RFamide neur
16	47	32.9	597	2	G82308	oxaloacetate decar
17	47	32.9	967	2	F82668	oxoglutarate dehyd
18	47	32.9	5138	2	B96695	hypothetical prote
19	46.5	32.5	314	2	T02961	annexin P33 - maiz
20	46.5	32.5	435	2	A44308	Antho-RFamide prec
21	46.5	32.5	1140	2	T09486	hypothetical prote
22	46	32.2	165	2	S59899	chlorocruorin chai
23	46	32.2	399	2	T35440	probable polyamine
24	46	32.2	946	1	IYHU2	inter-alpha-trypsi
25	46	32.2	1164	2	T24806	hypothetical prote
26	46	32.2	1378	2	A81393	DNA-directed RNA p
27	45.5	31.8	261	2	G69510	conserved hypothet
28	45.5	31.8	287	2	S43852	neuropeptide Pol-R
29	45.5	31.8	327	2	AF2859	conserved hypothet
30	45.5	31.8	327	2	D97636	probable secreted
31	45.5	31.8	562	2	C71473	hypothetical prote
32	45.5	31.8	905	2	G83314	NADH dehydrogenase
33	45	31.5	273	2	S06736	photosystem II oxy
34	45	31.5	273	2	AG2287	manganese-stabilzi
35	45	31.5	295	2	F83201	conserved hypothet
36	45	31.5	346	2	H95406	conserved hypothet
37	45	31.5	591	2	B44465	sodium ion pump ox
38	45	31.5	591	2	AB0509	oxaloacetate decar
39	45	31.5	591	2	AE0909	oxaloacetate decar
40	45	31.5	596	2	A28088	oxaloacetate decar
41	45	31.5	715	2	S52675	probable membrane
42	45	31.5	864	1	VCLJG4	env polyprotein -
43	45	31.5	915	2	B59433	chromosome 5 GAP-l
44	45	31.5	1199	2	T23005	hypothetical prote
45	45	31.5	1217	2	T22672	hypothetical prote

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 96.5%; Score 138; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 9.3e-13;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
||||||||||||||||||||||||||
Db 140 NLWAAQRYGRELRRMSDEFEGSFKGL 165

RESULT 2
JC5575
inter-alpha-trypsin inhibitor heavy chain 2 - golden hamster
C;Species: Mesocricetus auratus (golden hamster)
C;Date: 23-Sep-1997 #sequence_revision 23-Sep-1997 #text_change 20-Jun-2000
C;Accession: JC5575; PC4485
R;Nakatani, T.; Suzuki, Y.; Yamamoto, T.; Sinohara, H.
J. Biochem. 122, 71-82, 1997
A;Title: Molecular cloning and sequencing of cDNAs encoding three heavy-chain precurs
sin inhibitor heavy chain family.
A;Reference number: JC5574; MUID:97420688; PMID:9276673
A;Accession: JC5575
A;Molecule type: mRNA
A;Residues: 1-946 <NAK>
A;Cross-references: DDBJ:D89286; NID:g1694689; PIDN:BAA13939.1; PID:g1694690
A;Experimental source: liver
A;Accession: PC4485
A;Molecule type: protein
A;Residues: 55-64;140-146;151-156;424-447;500-528;577-605 <NA2>
C;Comment: In the plasma three inter-alpha-trypsin inhibitor heavy chains 1, 2 and 3
that the complexes play important role for panceatic cancer.
C;Superfamily: inter-alpha-trypsin inhibitor complex component II
F;261-264,717-916/Disulfide bonds: #status predicted

Query Match 37.8%; Score 54; DB 2; Length 946;
Best Local Similarity 34.6%; Pred. No. 9.9;

```
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHFQGV 237


RESULT 3
D70760
hypothetical protein Rv2014 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: D70760
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Sqares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: A70500; MUID:98295987; PMID:9634230
A;Accession: D70760
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-223 <COL>
A;Cross-references: GB:Z74025; GB:AL123456; NID:g3261586; PIDN:CAA98415.1; PID:e1299911;
A;Experimental source: strain H37Rv
C;Genetics:
A;Gene: Rv2014

  Query Match             37.1%;  Score 53;  DB 2;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 3.2;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 4
S54354
inter-alpha-inhibitor H2 chain - mouse
C;Species: Mus musculus (house mouse)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 20-Aug-1999
C;Accession: S54354
R;Chan, P.; Risler, J.L.; Raquenez, G.; Salier, J.P.
Biochem. J. 306, 505-512, 1995
A;Title: The three heavy-chain precursors for the inter-alpha-inhibitor family in mouse:
A;Reference number: S54353; MUID:95194326; PMID:7534067
A;Accession: S54354
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-946 <CHA>
A;Cross-references: EMBL:X70392; NID:g695633; PIDN:CAA49842.1; PID:g695634
C;Superfamily: inter-alpha-trypsin inhibitor complex component II

  Query Match             37.1%;  Score 53;  DB 2;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 14;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237


RESULT 5
S38185
2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) ARO4 - yeast (Saccharomyces ce
N;Alternate names: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DAHP synthase;
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C;Accession: S38185; S46126; S46130; JN0322; B48651
R;Doignon, F.; Biteau, N.; Aigle, M.; Crouzet, M.
Yeast 9, 1131-1137, 1993
A;Title: The complete sequence of a 6794 bp segment located on the right arm of chrom
A;Reference number: S38185; MUID:94078675; PMID:8256522
A;Accession: S38185
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <DOI>
A;Cross-references: GB:L20296; NID:g311101; PIDN:AAA65607.1; PID:g311102
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46126
A;Molecule type: DNA
A;Residues: 1-370 <ALJ>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46130
A;Molecule type: DNA
A;Residues: 1-370 <AIG>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Kuenzler, M.; Paravicini, G.; Egli, C.M.; Irniger, S.; Braus, G.H.
Gene 113, 67-74, 1992
A;Title: Cloning, primary structure and regulation of the ARO4 gene, encoding the tyr
A;Reference number: JN0322; MUID:92225349; PMID:1348717
A;Accession: JN0322
A;Molecule type: DNA
A;Residues: 1-204,208-370 <KUE>
A;Cross-references: EMBL:X61107
R;Kuenzler, M.; Balmelli, T.; Egli, C.M.; Paravicini, G.; Braus, G.H.
J. Bacteriol. 175, 5548-5558, 1993
A;Title: Cloning, primary structure, and regulation of the HIS7 gene encoding a bifun
A;Reference number: A48651; MUID:93374850; PMID:8366040
A;Accession: B48651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 352-370 <KU2>
A;Cross-references: GB:X61107
C;Comment: This enzyme catalyzes the condensation of phosphoenolpyruvate and D-erythr
C;Genetics:
A;Gene: SGD:ARO4
A;Cross-references: SGD:S0000453; MIPS:YBR249c
A;Map position: 2R
C;Function:
A;Description: aldehyde-lyase; carbon-carbon lyase
A;Pathway: aromatic amino acid biosynthesis; shikimate pathway
A;Note: first step in shikimate pathway
C;Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase
C;Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; cy

  Query Match             36.4%;  Score 52;  DB 2;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 7.5;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEG 22
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 6
A42095
floral homeotic protein APETALA3 (AP3) - Arabidopsis thaliana
N;Alternate names: homeotic protein APETALA3; MADS-box regulatory protein AP3
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C;Accession: A42095; S52633; T47593
R;Jack, T.; Brockman, L.L.; Meyerowitz, E.M.
Cell 68, 683-697, 1992
A;Title: The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS box and is
A;Reference number: A42095; MUID:92154682; PMID:1346756
A;Accession: A42095
A;Status: preliminary
A;Molecule type: mRNA
```

```
A;Residues: 1-232 <JAC>
A;Cross-references: GB:M86357; NID:g166607; PIDN:AAA32740.1; PID:g166608
A;Experimental source: petals, stamens
A;Note: sequence extracted from NCBI backbone (NCBIN:82520, NCBIP:82521)
R;Okamoto, H.; Yano, A.; Shiraishi, H.; Okada, K.; Shimura, Y.
Plant Mol. Biol. 26, 465-472, 1994
A;Title: Genetic complementation of a floral homeotic mutation, apetala3, with an Arabid
A;Reference number: S52633; MUID:95036018; PMID:7948893
A;Accession: S52633
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-63 <OKA>
A;Cross-references: GB:D21125
R;Bloecker, H.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quetier, F.; Salanoubat M.Mewes,
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24469
A;Accession: T47593
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-232 <BLO>
A;Cross-references: EMBL:AL132971
A;Experimental source: cultivar Columbia; BAC clone T12E18
C;Genetics:
A;Map position: 3
A;Introns: 63/2; 85/3; 106/2; 139/3; 153/3; 168/3
A;Note: T12E18.30
C;Superfamily: transcription factor squa; serum response factor DNA-binding domain homol
C;Keywords: DNA binding; nucleus; transcription regulation
F;2-57/Domain: serum response factor DNA-binding domain homology <SRF>

  Query Match             35.7%;  Score 51;  DB 2;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 6.5;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          7 QRYG--------RELRRMSDEFEGSFK 25
              || |        :|||:|| | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 7
C84338
spermidine/putrescine ABC transporter [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: C84338
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: C84338
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-374 <STO>
A;Cross-references: GB:AE004437; NID:g10581314; PIDN:AAG20071.1; GSPDB:GN00138
C;Genetics:
A;Gene: potA2

  Query Match             35.0%;  Score 50;  DB 2;  Length 374;
  Best Local Similarity   76.9%;  Pred. No. 15;
  Matches   10;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy         12 ELRRMSDEFEGSF 24
              ||||:||  ||||
Db        197 ELRRLSDAVEGSF 209


RESULT 8
A96753
probable threonine synthase [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C;Accession: A96753
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: A96753
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-516 <STO>
A;Cross-references: GB:AE005173; NID:g5903070; PIDN:AAD55628.1; GSPDB:GN00141
C;Genetics:
A;Gene: F3N23.1
A;Map position: 1

  Query Match             35.0%;  Score 50;  DB 2;  Length 516;
  Best Local Similarity   35.3%;  Pred. No. 21;
  Matches   12;  Conservative    7;  Mismatches    7;  Indels    8;  Gaps    1;

Qy          2 NLWAAQRYGRELRRMSD--------EFEGSFKGL 27
              ||: |:|:|::  :|:|          |||| |
Db        163 NLFWAERFGKQYLQMNDLWVKHCGISHTGSFKDL 196


RESULT 9
E83517
conserved hypothetical protein PA1031 [imported] - Pseudomonas aeruginosa (strain PAO
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C;Accession: E83517
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83517
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-453 <STO>
A;Cross-references: GB:AE004535; GB:AE004091; NID:g9946936; PIDN:AAG04420.1; GSPDB:GN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA1031

  Query Match             34.3%;  Score 49;  DB 2;  Length 453;
  Best Local Similarity   55.6%;  Pred. No. 25;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          4 WAAQRYGR--ELRRMSDE 19
              ||::| ||  ||||:: |
Db         65 WASERQGREEELRRLASE 82


RESULT 10
S40376
Ig kappa chain - human
C;Species: Homo sapiens (man)
C;Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C;Accession: S40376
R;Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3248-3271, 1993
A;Title: Expressed human immunoglobulin chi genes and their hypermutation.
A;Reference number: S40312; MUID:94080891; PMID:8258341
A;Accession: S40376
A;Status: preliminary; translation not shown
```

```
A;Molecule type: mRNA
A;Residues: 1-134 <KLE>
A;Cross-references: EMBL:X72486; NID:g441440; PIDN:CAA51154.1; PID:g441441
C;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin
F;34-113/Domain: immunoglobulin homology <IMM>

  Query Match             33.9%;  Score 48.5;  DB 2;  Length 134;
  Best Local Similarity   38.2%;  Pred. No. 8.6;
  Matches   13;  Conservative    1;  Mismatches    9;  Indels   11;  Gaps    1;

Qy          4 WAAQRYGRELRRM----------SDEFEGSFKG 26
              |  || ||  ||:           || | ||  |
Db         58 WFRQRPGRSPRRLIYNVSKRDSGVSDRFSGSGSG 91


RESULT 11
T02975
annexin P35 - maize
C;Species: Zea mays (maize)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 21-Jul-2000
C;Accession: T02975
R;Battey, N.H.; James, N.C.; Greenland, A.J.
Plant Physiol. 112, 1391-1396, 1996
A;Title: CDNA isolation and gene expression of maize annexins P33 and P35.
A;Reference number: Z14796; MUID:97092863; PMID:8938425
A;Accession: T02975
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-314 <BAT>
A;Cross-references: EMBL:X98245; NID:g1370602; PIDN:CAA66901.1; PID:g1370603
A;Experimental source: cultivar clipper; root tip
C;Superfamily: annexin I; annexin repeat homology
F;14-85/Domain: annexin repeat homology <AXR>

  Query Match             33.9%;  Score 48.5;  DB 2;  Length 314;
  Best Local Similarity   47.6%;  Pred. No. 21;
  Matches   10;  Conservative    4;  Mismatches    6;  Indels    1;  Gaps    1;

Qy          6 AQRYGRE-LRRMSDEFEGSFK 25
              |: ||:| || : ||  | |:
Db         54 AEAYGKELLRALGDEIHGKFE 74


RESULT 12
C36365
transforming protein homolog MRAS3 - Rhizomucor racemosus
C;Species: Rhizomucor racemosus
C;Date: 28-Mar-1991 #sequence_revision 28-Mar-1991 #text_change 19-Jan-2001
C;Accession: C36365
R;Casale, W.L.; McConnell, D.G.; Wang, S.Y.; Lee, Y.J.; Linz, J.E.
Mol. Cell. Biol. 10, 6654-6663, 1990
A;Title: Expression of a gene family in the dimorphic fungus Mucor racemosus which exhib
A;Reference number: A36365; MUID:91061774; PMID:1701021
A;Accession: C36365
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-206 <CAS>
A;Cross-references: GB:M55177
C;Superfamily: ras transforming protein; translation elongation factor Tu homology
C;Keywords: GTP binding; nucleotide binding; P-loop
F;11-126/Domain: translation elongation factor Tu homology <ETU>
F;17-24/Region: nucleotide-binding motif A (P-loop)
F;123-126/Region: GTP-binding NKXD motif
F;153-155/Region: GTP-binding SAK/L motif
F;23,24,42,123,124,126,153/Binding site: Mg-GTP (Lys, Ser, Thr, Asn, Lys, Asp, Ser) #sta

  Query Match             33.6%;  Score 48;  DB 2;  Length 206;
  Best Local Similarity   62.5%;  Pred. No. 16;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         11 RELRRMSDEFEGSFKG 26
              ||:|||: | ||  ||
Db        169 REIRRMNKEQEGRSKG 184


RESULT 13
F72289
oxidoreductase, sol/devB family - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72289
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hic
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <ARN>
A;Cross-references: GB:AE001772; GB:AE000512; NID:g4981693; PIDN:AAD36230.1; PID:g498
A;Experimental source: strain MSB8
C;Genetics:
A;Gene: TM1154
C;Superfamily: yeast SOL3 protein

  Query Match             33.6%;  Score 48;  DB 2;  Length 220;
  Best Local Similarity   34.8%;  Pred. No. 17;
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AAQRYGRELRRMSDEFEGSFKGL 27
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133


RESULT 14
T08545
threonine synthase (EC 4.2.3.1) precursor - Arabidopsis thaliana
N;Alternate names: protein F27B13.80
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Jun-2002
C;Accession: T08545; S71362; S74307
R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Bancroft, I.; Mewes, H.W.;
submitted to the Protein Sequence Database, May 1999
A;Reference number: Z16442
A;Accession: T08545
A;Molecule type: DNA
A;Residues: 1-526 <BEV>
A;Cross-references: EMBL:AL050352; GSPDB:GN00062; ATSP:F27B13.80
A;Experimental source: cultivar Columbia; BAC clone F27B13
R;Curien, G.; Dumas, R.; Ravanel, S.; Douce, R.
FEBS Lett. 390, 85-90, 1996
A;Title: Characterization of an Arabidopsis thaliana cDNA encoding an S-adenosylmethi
A;Reference number: S71362; MUID:96314555; PMID:8706836
A;Accession: S71362
A;Molecule type: mRNA
A;Residues: 'L',3-526 <CUR>
A;Cross-references: EMBL:L41666; NID:g1448916; PIDN:AAB04607.1; PID:g1448917
A;Accession: S74307
A;Molecule type: protein
A;Residues: 40-54 <CUI>
C;Genetics:
A;Gene: ATSP:F27B13.80
A;Map position: 4
A;Genome: nuclear
C;Keywords: carbon-oxygen lyase; chloroplast
F;1-39/Domain: transit peptide (chloroplast) #status predicted <TNP>
F;40-526/Product: threonine synthase #status experimental <MAT>

  Query Match             33.6%;  Score 48;  DB 2;  Length 526;
  Best Local Similarity   35.3%;  Pred. No. 42;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    8;  Gaps    1;

Qy          2 NLWAAQRYGRELRRMSD--------EFEGSFKGL 27
```

```
              ||: |:|:|::   |:|          |||| |
Db        172 NLFWAERFGKQFLGMNDLWVKHCGISHTGSFKDL 205


RESULT 15
A39172
Antho-RFamide neuropeptide 19 repeat precursor - sea anemone (Calliactis parasitica)
C;Species: Calliactis parasitica
C;Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 21-Jul-2000
C;Accession: A39172
R;Darmer, D.; Schmutzler, C.; Diekhoff, D.; Grimmelikhuijzen, C.J.P.
Proc. Natl. Acad. Sci. U.S.A. 88, 2555-2559, 1991
A;Title: Primary structure of the precursor for the sea anemone neuropeptide Antho-RFami
A;Reference number: A39172; MUID:91172845; PMID:1706527
A;Accession: A39172
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-334 <DAR>
A;Cross-references: GB:M59166; NID:g156133; PIDN:AAA27878.1; PID:g156134
C;Keywords: neuropeptide

  Query Match             33.2%;  Score 47.5;  DB 2;  Length 334;
  Best Local Similarity   44.0%;  Pred. No. 31;
  Matches   11;  Conservative    3;  Mismatches   10;  Indels    1;  Gaps    1;

Qy          1 KNLWAAQRYGRELR-RMSDEFEGSF 24
              |  :   ||||| : |   ||:| |
Db         89 KRRYVPGRYGREFQGRFGREFQGRF 113


Search completed: September 15, 2003, 17:27:04
Job time : 13.15 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 6.36429 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	138	96.5	204	1	BAD_MOUSE	Q61337 mus musculu
2	138	96.5	205	1	BAD_RAT	O35147 rattus norv
3	114	79.7	168	1	BAD_HUMAN	Q92934 homo sapien
4	54	37.8	946	1	ITH2_MESAU	P97279 mesocricetu
5	53	37.1	946	1	ITH2_MOUSE	Q61703 mus musculu
6	52.5	36.7	506	1	MATK_LEDPA	O62992 ledum palus
7	52.5	36.7	506	1	MATK_RHOFR	O62984 rhododendro
8	52.5	36.7	506	1	MATK_RHOTS	O62991 rhododendro
9	52	36.4	370	1	AROG_YEAST	P32449 saccharomyc
10	51	35.7	232	1	AP3_ARATH	P35632 arabidopsis
11	51	35.7	851	1	CE05_MOUSE	Q8k2h3 mus musculu
12	49	34.3	453	1	RMUC_PSEAE	Q9i4u3 pseudomonas
13	48	33.6	205	1	RAS3_RHIRA	P22280 rhizomucor
14	48	33.6	220	1	6PGL_THEMA	Q9x0n8 thermotoga
15	48	33.6	519	1	THRC_SOLTU	Q9mt28 solanum tub
16	48	33.6	526	1	THRC_ARATH	Q9s7b5 arabidopsis
17	47.5	33.2	334	1	FMRA_CALPA	Q01133 calliactis
18	47.5	33.2	507	1	MATK_LOIPR	O47169 loiseleuria
19	47	32.9	198	1	BIM_HUMAN	O43521 homo sapien
20	46.5	32.5	429	1	FMR2_ANTEL	Q16994 anthopleura
21	46.5	32.5	435	1	FMR1_ANTEL	P10419 anthopleura
22	46	32.2	946	1	ITH2_HUMAN	P19823 homo sapien
23	46	32.2	1378	1	RPOB_CAMJE	Q46124 campylobact
24	45.5	31.8	287	1	PRFA_POLPE	P21259 polyorchis
25	45	31.5	273	1	PSBO_ANASP	P13907 anabaena sp
26	45	31.5	328	1	SNF4_KLULA	Q9p869 kluyveromyc
27	45	31.5	590	1	DCOA_SALTY	Q03030 salmonella
28	45	31.5	595	1	DCOA_KLEPN	P13187 klebsiella
29	45	31.5	653	1	HT2A_HUMAN	Q13049 homo sapien
30	45	31.5	865	1	ENV_SIVAT	P05886 simian immu
31	45	31.5	915	1	CE05_HUMAN	Q9nyf5 homo sapien
32	45	31.5	1535	1	LML1_CAEEL	Q18823 caenorhabdi
33	44.5	31.1	506	1	MATK_GAUPR	Q95gj0 gaultheria
34	44.5	31.1	512	1	MATK_LILTS	Q9gig3 lilium tsin
35	44.5	31.1	907	1	NUOG_ECOLI	P33602 escherichia
36	44.5	31.1	907	1	NUOG_SALTY	P33900 salmonella
37	44	30.8	196	1	BIM_MOUSE	O54918 mus musculu
38	44	30.8	196	1	BIM_RAT	O88498 rattus norv
39	44	30.8	262	1	END8_ECO57	Q8x9c6 escherichia
40	44	30.8	262	1	END8_ECOLI	P50465 escherichia
41	44	30.8	262	1	END8_SALTI	Q8z8d2 salmonella
42	44	30.8	262	1	END8_SALTY	Q8zqu6 salmonella
43	44	30.8	629	1	SYM_THEMA	O33925 thermotoga
44	44	30.8	768	1	ENV_SIVA1	P27757 simian immu
45	44	30.8	877	1	ENV_SIVAG	P27977 simian immu

ALIGNMENTS

RESULT 1
BAD_MOUSE
ID BAD_MOUSE STANDARD; PRT; 204 AA.
AC Q61337;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE 6) (Bcl-xL/Bcl-2 associated death promoter).
GN BAD OR BBC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Thymus;
RX MEDLINE=95136361; PubMed=7834748;
RA Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT promotes cell death.";
RL Cell 80:285-291(1995).
RN [2]
RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX MEDLINE=98022383; PubMed=9381178;
RA Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT "Interleukin-3-induced phosphorylation of BAD through the protein
RT kinase Akt.";
RL Science 278:687-689(1997).
RN [3]
RP MUTAGENESIS OF SERINE RESIDUES.
RX MEDLINE=20403302; PubMed=10949026;
RA Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA Greenberg M.E.;
RT "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT BH3 domain phosphorylation.";
RL Mol. Cell 6:41-51(2000).
CC -!- FUNCTION: Promotes cell death. Successfully competes for the
CC binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC of heterodimerization of these proteins with BAX. Can reverse the
CC death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC Appears to act as a link between growth factor receptor signaling
CC and the apoptotic pathways.
CC -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC phosphorylation, locates to the cytoplasm.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC with 14-3-3 proteins. This interaction then facilitates the
CC phosphorylation at Ser-155, a site within the BH3 domain, leading
CC to the release of Bcl-x(L) and the promotion of cell survival.

```
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             96.5%;  Score 138;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        140 NLWAAQRYGRELRRMSDEFEGSFKGL 165


RESULT 2
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMPPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;
```

```
  Query Match             96.5%;  Score 138;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-13;
  Matches   26;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              ||||||||||||||||||||||||||
Db        141 NLWAAQRYGRELRRMSDEFEGSFKGL 166
```

```
RESULT 3
BAD_HUMAN
ID   BAD_HUMAN      STANDARD;      PRT;   168 AA.
AC   Q92934; O14803;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE   XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN   BAD OR BBC6 OR BCL2L8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT   "A human protein that interacts with Bcl-2 and have homology to mouse
RT   BAD.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX   MEDLINE=97083574; PubMed=8929532;
RA   Wang H.-G., Rapp U.R., Reed J.C.;
RT   "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL   Cell 87:629-638(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Takayama S., Reed J.C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND DIMERIZATION.
RC   TISSUE=Bone marrow;
RX   MEDLINE=98049554; PubMed=9388232;
RA   Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA   Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT   "Dimerization properties of human BAD.";
RL   J. Biol. Chem. 272:30866-30872(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are ho restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;
```

```
  Query Match             79.7%;  Score 114;  DB 1;  Length 168;
```

```
  Best Local Similarity   91.7%;  Pred. No. 5.8e-10;
  Matches   22;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSPK 25
              |||||||||||||||||||  |||
Db        103 NLWAAQRYGRELRRMSDEFVDSFK 126


RESULT 4
ITH2_MESAU
ID   ITH2_MESAU     STANDARD;      PRT;   946 AA.
AC   P97279;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2) (HC2).
GN   ITIH2.
OS   Mesocricetus auratus (Golden hamster).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC   Mesocricetus.
OX   NCBI_TaxID=10036;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RX   MEDLINE=97420688; PubMed=9276673;
RA   Nakatani T., Suzuki Y., Yamamoto T., Sinohara H.;
RT   "Molecular cloning and sequencing of cDNAs encoding three heavy-chain
RT   precursors of the inter-alpha-trypsin inhibitor in Syrian hamster:
RT   implications for the evolution of the inter-alpha-trypsin inhibitor
RT   heavy chain family.";
RL   J. Biochem. 122:71-82(1997).
RN   [2]
RP   SEQUENCE OF 55-64; 140-146; 151-156; 424-447; 500-528 AND 577-605,
RP   AND SUBUNITS.
RC   TISSUE=Plasma;
RX   MEDLINE=97018241; PubMed=8864857;
RA   Yamamoto T., Yamamoto K., Sinohara H.;
RT   "Inter-alpha-trypsin inhibitor and its related proteins in Syrian
RT   hamster urine and plasma.";
RL   J. Biochem. 120:145-152(1996).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D89286; BAA13939.1; -.
DR   PIR; JC5575; JC5575.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                 H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    578    578       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                 (BY SIMILARITY).
FT   CONFLICT    510    510       V -> Y (IN REF. 2).
FT   CONFLICT    595    595       E -> I (IN REF. 2).
SQ   SEQUENCE   946 AA;  106580 MW;  CA8BF565458E7B2E CRC64;

  Query Match             37.8%;  Score 54;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 3.7;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSPKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIVELQGMRFLHVPDTFEGHPQGV 237


RESULT 5
ITH2_MOUSE
ID   ITH2_MOUSE     STANDARD;      PRT;   946 AA.
AC   Q61703;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Inter-alpha-trypsin inhibitor heavy chain H2 precursor (ITI heavy
DE   chain H2) (Inter-alpha-inhibitor heavy chain 2).
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6N; TISSUE=Liver;
RX   MEDLINE=95194326; PubMed=7534067;
RA   Chan P., Risler J.-L., Raguenez G., Salier J.-P.;
RT   "The three heavy-chain precursors for the inter-alpha-inhibitor
RT   family in mouse: new members of the multicopper oxidase protein group
RT   with differential transcription in liver and brain.";
RL   Biochem. J. 306:505-512(1995).
CC   -!- FUNCTION: MAY ACT AS A CARRIER OF HYALURONAN IN SERUM OR AS A
CC       BINDING PROTEIN BETWEEN HYALURONAN AND OTHER MATRIX PROTEIN,
CC       INCLUDING THOSE ON CELL SURFACES IN TISSUES TO REGULATE THE
CC       LOCALIZATION, SYNTHESIS AND DEGRADATION OF HYALURONAN WHICH ARE
CC       ESSENTIAL TO CELLS UNDERGOING BIOLOGICAL PROCESSES (BY
CC       SIMILARITY).
CC   -!- SUBUNIT: I-ALPHA-I PLASMA PROTEASE INHIBITORS ARE ASSEMBLED FROM
CC       ONE OR TWO HEAVY CHAINS (H1, H2 OR H3) AND ONE LIGHT CHAIN,
CC       BIKUNIN. INTER-ALPHA-INHIBITOR (I-ALPHA-I) IS COMPOSED OF H1, H2
CC       AND BIKUNIN, INTER-ALPHA-LIKE INHIBITOR (I-ALPHA-LI) OF H2 AND
CC       BIKUNIN, AND PRE-ALPHA-INHIBITOR (P-ALPHA-I) OF H3 AND BIKUNIN.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN BOTH LIVER AND BRAIN.
CC   -!- PTM: HEAVY CHAINS ARE INTERLINKED WITH BIKUNIN VIA A CHONDROITIN
CC       4-SULFATE BRIDGE TO THE THEIR C-TERMINAL ASPARTATE (BY
CC       SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO THE ITIH FAMILY.
CC   -!- SIMILARITY: Contains 1 VWFA domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
```

```
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X70392; CAA49842.1; -.
DR   PIR; S54354; S54354.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   Pfam; PF00092; vwa; 1.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
KW   Serine protease inhibitor; Repeat; Signal; Multigene family;
KW   Glycoprotein.
FT   SIGNAL        1     18       POTENTIAL.
FT   PROPEP       19     54       BY SIMILARITY.
FT   CHAIN        55    702       INTER-ALPHA-TRYPSIN INHIBITOR HEAVY CHAIN
FT                                H2.
FT   PROPEP      703    946       BY SIMILARITY.
FT   DOMAIN      308    468       VWFA.
FT   CARBOHYD    118    118       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    263    263       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    445    445       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   BINDING     702    702       CHONDROITIN 4-SULFATE, CROSS-LINK SITE
FT                                (BY SIMILARITY).
SQ   SEQUENCE   946 AA;  105927 MW;  40DB6716433ED9DC CRC64;

  Query Match             37.1%;  Score 53;  DB 1;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 5.2;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIIEPQGMRFLHVPDTFEGHFQGV 237


RESULT 6
MATK_LEDPA
ID   MATK_LEDPA     STANDARD;      PRT;   506 AA.
AC   O62992;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Ledum palustre (Wild rosemary).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75583;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012751; BAA25872.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60412 MW;  CFEA926307DAC85E CRC64;

  Query Match             36.7%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 3.1;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 7
MATK_RHOFR
ID   MATK_RHOFR     STANDARD;      PRT;   506 AA.
AC   O62984;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron ferrugineum (Alpenrose).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49622;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012741; BAA25862.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60534 MW;  ADA44B25E92436E8 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 3.1;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 8
MATK_RHOTS
ID   MATK_RHOTS     STANDARD;      PRT;   506 AA.
AC   O62991;
```

```
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Maturase K (Intron maturase).
GN   MATK.
OS   Rhododendron tsusiophyllum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49629;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of sectional relationships in the genus
RT   Rhododendron (Ericaceae) based on matK sequences.";
RL   Shokubutsu Kenkyu Zasshi 73:143-154(1998).
CC   -!- FUNCTION: Probably assists in splicing chloroplast group II
CC       introns (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE INTRON MATURASE FAMILY 2. MATK
CC       SUBFAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AB012750; BAA25871.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60569 MW;  AEE12FF8809C223E CRC64;

  Query Match             36.7%;  Score 52.5;  DB 1;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 3.1;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 9
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match             36.4%;  Score 52;  DB 1;  Length 370;
  Best Local Similarity   47.6%;  Pred. No. 2.7;
  Matches   10;  Conservative    5;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEG 22
              :| ||| |   |:::||| :|
Db         80 DLEAAQEYALRLKKLSDELKG 100


RESULT 10
AP3_ARATH
ID   AP3_ARATH      STANDARD;      PRT;   232 AA.
AC   P35632; Q39003; Q8LB79; Q9S7Q3; Q9SQ14; Q9SQ15; Q9SQ16; Q9SQ17;
AC   Q9SQ18; Q9SQ19; Q9SQ20; Q9SQ21; Q9SQ22; Q9SX13;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Floral homeotic protein APETALA3.
GN   AP3 OR AT3G54340 OR T12E18_30.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Petal;
RX MEDLINE=92154682; PubMed=1346756;
RA Jack T., Brockman L.L., Meyerowitz E.M.;
RT "The homeotic gene APETALA3 of Arabidopsis thaliana encodes a MADS
RT box and is expressed in petals and stamens.";
RL Cell 68:683-697(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Landsberg erecta;
RX MEDLINE=95036018; PubMed=7948893;
RA Okamoto H., Yano A., Shiraishi H., Okada K., Shimura Y.;
RT "Genetic complementation of a floral homeotic mutation, apetala3,
RT with an Arabidopsis thaliana gene homologous to DEFICIENS of
RT Antirrhinum majus.";
RL Plant Mol. Biol. 26:465-472(1994).
RN [3]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC STRAIN=cv. Bla-1, cv. Bretagny, cv. Bs-1, cv. Bu-0, cv. Bu-2,
RC cv. Chi-1, cv. Co-1, cv. Columbia, cv. Corsacalla-1, cv. Cvi-0,
RC cv. Gr-3, cv. Jl-1, cv. Kas-1, cv. Kent, cv. Landsberg erecta,
RC cv. Li-3, cv. Li-8, and cv. Lisse;
RX MEDLINE=99126449; PubMed=9927474;
RA Purugganan M.D., Suddith J.I.;
RT "Molecular population genetics of floral homeotic loci: departures
RT from the equilibrium-neutral model at the APETALA3 and PISTILLATA
RT genes of Arabidopsis thaliana.";
RL Genetics 151:839-848(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=21016720; PubMed=11130713;
RA Salanoubat M., Lemcke K., Rieger M., Ansorge W., Unseld M.,
RA Fartmann B., Valle G., Bloecker H., Perez-Alonso M., Obermaier B.,
RA Delseny M., Boutry M., Grivell L.A., Mache R., Puigdomenech P.,
RA De Simone V., Choisne N., Artiguenave F., Robert C., Brottier P.,
RA Wincker P., Cattolico L., Weissenbach J., Saurin W., Quetier F.,
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Benes V.,
RA Wurmbach E., Drzonek H., Erfle H., Jordan N., Bangert S.,
RA Wiedelmann R., Kranz H., Voss H., Holland R., Brandt P., Nyakatura G.,
RA Vezzi A., D'Angelo M., Pallavicini A., Toppo S., Simionati B.,
RA Conrad A., Hornischer K., Kauer G., Loehnert T.-H., Nordsiek G.,
RA Reichelt J., Scharfe M., Schoen O., Bargues M., Terol J., Climent J.,
RA Navarro P., Collado C., Perez-Perez A., Ottenwaelder B., Duchemin D.,
RA Cooke R., Laudie M., Berger-Llauro C., Purnelle B., Masuy D.,
RA de Haan M., Maarse A.C., Alcaraz J.-P., Cottet A., Casacuberta E.,
RA Monfort A., Argiriou A., Flores M., Liguori R., Vitale D.,
RA Mannhaupt G., Haase D., Schoof H., Rudd S., Zaccaria P., Mewes H.-W.,
RA Mayer K.F.X., Kaul S., Town C.D., Koo H.L., Tallon L.J., Jenkins J.,
RA Rooney T., Rizzo M., Walts A., Utterback T., Fujii C.Y., Shea T.P.,
RA Creasy T.H., Haas B., Maiti R., Wu D., Peterson J., Van Aken S.,
RA Pai G., Militscher J., Sellers P., Gill J.E., Feldblyum T.V.,
RA Preuss D., Lin X., Nierman W.C., Salzberg S.L., White O., Venter J.C.,
RA Fraser C.M., Kaneko T., Nakamura Y., Sato S., Kato T., Asamizu E.,
RA Sasamoto S., Kimura T., Idesawa K., Kawashima K., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakayama S., Nakazaki N., Shinpo S., Takeuchi C., Wada T.,
RA Watanabe A., Yamada M., Yasuda M., Tabata S.;
RT "Sequence and analysis of chromosome 3 of the plant Arabidopsis
RT thaliana.";
RL Nature 408:820-822(2000).
RN [5]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RT "Full-length cDNA from Arabidopsis thaliana.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;
RT "RIKEN Arabidopsis full length cDNA clones (RAFLs) sequenced by the
RT SSP consortium (Salk/Stanford/PGEC).";
RL Submitted (AUG-2002) to the EMBL/GenBank/DDBJ databases.
RN [7]
RP SEQUENCE OF 36-128 FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=99311297; PubMed=10382288;
RA Brunel D., Froger N., Pelletier G.;
RT "Development of amplified consensus genetic markers (ACGM) in Brassica
RT napus from Arabidopsis thaliana sequences of known biological
RT function.";
RL Genome 42:387-402(1999).
RN [8]
RP FUNCTION.
RX PubMed=8565821;
RA Krizek B.A., Meyerowitz E.M.;
RT "The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient
RT to provide the B class organ identity function.";
RL Development 122:11-22(1996).
RN [9]
RP CHARACTERIZATION.
RX PubMed=8643482;
RA Riechmann J.L., Krizek B.A., Meyerowitz E.M.;
RT "Dimerization specificity of Arabidopsis MADS domain homeotic proteins
RT APETALA1, APETALA3, PISTILLATA, and AGAMOUS.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:4793-4798(1996).
RN [10]
RP GENETIC REGULATION.
RX PubMed=11283333;
RA Ng M., Yanofsky M.F.;
RT "Activation of the Arabidopsis B class homeotic genes by APETALA1.";
RL Plant Cell 13:739-753(2001).
RN [11]
RP CHARACTERIZATION.
RX PubMed=11206550;
RA Honma T., Goto K.;
RT "Complexes of MADS-box proteins are sufficient to convert leaves into
RT floral organs.";
RL Nature 409:525-529(2001).
CC -!- FUNCTION: Probable transcription factor involved in the genetic
CC control of flower development. Is required for normal development
CC of petals and stamens in the wild-type flower. Forms an
CC heterodimer with PISTILLATA that is required for autoregulation of
CC both AP3 and PI genes. AP3/PI heterodimer interacts with APETALA1
CC or SEPALLATA3 to form a ternary complex that could be responsible
CC for the regulation of the genes involved in the flower
CC development.
CC -!- SUBUNIT: Forms an heterodimer with PISTILLATA, capable of binding
CC to CArG-box sequences. AP3/PI heterodimer binds AP1 or SEP3 to
CC form complexes.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Expressed in petals and stamens.
CC -!- INDUCTION: Positively regulated by the meristem identity proteins
CC APETALA1 and LEAFY with the cooperation of UFO.
CC -!- MISCELLANEOUS: Mutations in AP3 cause transformation of petals
CC into sepals and stamina into carpels.
CC -!- SIMILARITY: BELONGS TO THE MADS DOMAIN FAMILY OF TRANSCRIPTION
CC FACTORS.
CC -!- SIMILARITY: Contains 1 K-box dimerization domain.
CC ---

CC ---
DR EMBL; M86357; AAA32740.1; -.
DR EMBL; D21125; BAA04665.1; -.
DR EMBL; AF115798; AAD51887.1; -.
DR EMBL; AF115799; AAD51888.1; -.

```
DR   EMBL; AF115800; AAD51889.1; -.
DR   EMBL; AF115801; AAD51890.1; -.
DR   EMBL; AF115802; AAD51891.1; -.
DR   EMBL; AF115803; AAD51892.1; -.
DR   EMBL; AF115804; AAD51893.1; -.
DR   EMBL; AF115805; AAD51894.1; -.
DR   EMBL; AF115806; AAD51895.1; -.
DR   EMBL; AF115807; AAD51896.1; -.
DR   EMBL; AF115808; AAD51897.1; -.
DR   EMBL; AF115809; AAD51898.1; -.
DR   EMBL; AF115810; AAD51899.1; -.
DR   EMBL; AF115811; AAD51900.1; -.
DR   EMBL; AF115812; AAD51901.1; -.
DR   EMBL; AF115813; AAD51902.1; -.
DR   EMBL; AF115814; AAD51903.1; -.
DR   EMBL; AL132971; CAB81799.1; -.
DR   EMBL; AY087369; AAM64919.1; -.
DR   EMBL; AY070397; AAL49893.1; -.
DR   EMBL; AY142590; AAN13159.1; -.
DR   EMBL; AF056541; AAD41557.1; -.
DR   PIR; A42095; A42095.
DR   HSSP; P11746; 1MNM.
DR   TRANSFAC; T01776; -.
DR   InterPro; IPR002487; TF_Kbox.
DR   InterPro; IPR002100; TF_MADSbox.
DR   Pfam; PF01486; K-box; 1.
DR   Pfam; PF00319; SRF-TF; 1.
DR   PRINTS; PR00404; MADSDOMAIN.
DR   SMART; SM00432; MADS; 1.
DR   PROSITE; PS00350; MADS_BOX_1; 1.
DR   PROSITE; PS50066; MADS_BOX_2; 1.
KW   Flowering; Transcription regulation; Activator; Developmental protein;
KW   Nuclear protein; DNA-binding; Coiled coil; Polymorphism.
FT   DOMAIN        3     57       MADS.
FT   DOMAIN       93    165       K-BOX.
FT   DOMAIN       75    164       COILED COIL (POTENTIAL).
FT   VARIANT      31     31       K -> R (in strain cv. Lisse).
FT   VARIANT      47     47       M -> T (in strain cv. Bretagny).
FT   VARIANT      61     61       N -> D (in strain cv. Corsacalla-1).
FT   VARIANT      73     73       T -> S (in strain cv. Li-8).
FT   VARIANT     109    109       L -> V (in strain cv. Kas-1).
FT   VARIANT     115    115       E -> K (in strains cv. Chi-1 and cv.
FT                                Gr-3).

  Query Match             35.7%;  Score 51;  DB 1;  Length 232;
  Best Local Similarity   44.4%;  Pred. No. 2.3;
  Matches   12;  Conservative    3;  Mismatches    4;  Indels    8;  Gaps    1;

Qy          7 QRYG--------RELRRMSDEFEGSFK 25
              || |        :||||: || | :||
Db        107 QRLGECLDELDIQELRRLEDEMENTFK 133


RESULT 11
CE05_MOUSE
ID   CE05_MOUSE     STANDARD;      PRT;   851 AA.
AC   Q8K2H3;
DT   15-SEP-2003 (Rel. 42, Created)
DT   15-SEP-2003 (Rel. 42, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Protein C5orf5 homolog.
GN   C5ORF5.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FVB/N;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SIMILARITY: Belongs to the FAM13 family.
CC   -!- SIMILARITY: Contains 1 Rho-GAP domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; BC031465; AAH31465.1; -.
DR   InterPro; IPR000198; RhoGAP.
DR   Pfam; PF00620; RhoGAP; 1.
DR   SMART; SM00324; RhoGAP; 1.
DR   PROSITE; PS50238; RHOGAP; 1.
KW   GTPase activation.
FT   DOMAIN       23    212       RHO-GAP.
FT   DOMAIN      189    256       GLU-RICH.
SQ   SEQUENCE   851 AA;  97054 MW;  C2B26669FB6DB2CE CRC64;

  Query Match             35.7%;  Score 51;  DB 1;  Length 851;
  Best Local Similarity   41.7%;  Pred. No. 9.3;
  Matches   10;  Conservative    6;  Mismatches    8;  Indels    0;  Gaps    0;

Qy          1 KNLWAAQRYGRELRRMSDEFEGSF 24
              : || |:   ::||:| |||  :|
Db        782 EQLWKARAEKKKLRKMLREFEEAF 805


RESULT 12
RMUC_PSEAE
ID   RMUC_PSEAE     STANDARD;      PRT;   453 AA.
AC   Q9I4U3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA recombination protein rmuC homolog.
GN   RMUC OR PA1031.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
```

```
CC   -!- FUNCTION: Involved in DNA recombination (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE RMUC FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE004535; AAG04420.1; -.
DR   PIR; E83517; E83517.
DR   InterPro; IPR003798; DUF195.
DR   Pfam; PF02646; RmuC; 1.
KW   DNA recombination; Coiled coil; Complete proteome.
FT   DOMAIN       16    201       COILED COIL (POTENTIAL).
SQ   SEQUENCE   453 AA;  51539 MW;  1E7EA97E82EC5E4B CRC64;

  Query Match             34.3%;  Score 49;  DB 1;  Length 453;
  Best Local Similarity   55.6%;  Pred. No. 9.3;
  Matches   10;  Conservative    4;  Mismatches    2;  Indels    2;  Gaps    1;

Qy          4 WAAQRYGR--ELRRMSDE 19
              ||::| ||  ||||:: |
Db         65 WASERQGREEELRRLASE 82


RESULT 13
RAS3_RHIRA
ID   RAS3_RHIRA     STANDARD;      PRT;   205 AA.
AC   P22280;
DT   01-AUG-1991 (Rel. 19, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Ras-like protein 3.
GN   RAS3.
OS   Rhizomucor racemosus (Mucor circinelloides f. lusitanicus).
OC   Eukaryota; Fungi; Zygomycota; Zygomycetes; Mucorales; Mucoraceae;
OC   Mucor.
OX   NCBI_TaxID=4841;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 1216B;
RX   MEDLINE=91061774; PubMed=1701021;
RA   Casale W.L., McConnell D.G., Wang S.-Y., Lee Y.-J., Linz J.E.;
RT   "Expression of a gene family in the dimorphic fungus Mucor racemosus
RT   which exhibits striking similarity to human ras genes.";
RL   Mol. Cell. Biol. 10:6654-6663(1990).
CC   -!- ENZYME REGULATION: ALTERNATE BETWEEN AN INACTIVE FORM BOUND TO GDP
CC       AND AN ACTIVE FORM BOUND TO GTP. ACTIVATED BY A GUANINE
CC       NUCLEOTIDE-EXCHANGE FACTOR (GEF) AND INACTIVATED BY A GTPASE-
CC       ACTIVATING PROTEIN (GAP).
CC   -!- SUBCELLULAR LOCATION: Plasma membrane.
CC   -!- DEVELOPMENTAL STAGE: IN SPORULATING MYCELIUM AND MUCH LESS IN
CC       GERMLING AND YEAST.
CC   -!- SIMILARITY: BELONGS TO THE SMALL GTPASE SUPERFAMILY. RAS FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M55177; AAA83379.1; -.
DR   PIR; C36365; C36365.
DR   HSSP; P01112; 1PLL.
DR   InterPro; IPR003577; GTPase_Ras.
DR   InterPro; IPR001806; Ras_trnsfrmng.
DR   InterPro; IPR005225; Small_GTP.
DR   Pfam; PF00071; ras; 1.
DR   PRINTS; PR00449; RASTRNSFRMNG.
DR   SMART; SM00173; RAS; 1.
DR   TIGRFAMs; TIGR00231; small_GTP; 1.
KW   GTP-binding; Prenylation; Lipoprotein.
FT   NP_BIND      16     23       GTP (BY SIMILARITY).
FT   NP_BIND      63     67       GTP (BY SIMILARITY).
FT   NP_BIND     122    125       GTP (BY SIMILARITY).
FT   DOMAIN       38     46       EFFECTOR REGION (PROBABLE).
FT   LIPID       202    202       FARNESYL (BY SIMILARITY).
SQ   SEQUENCE   205 AA;  23408 MW;  DBF086466F090F50 CRC64;

  Query Match             33.6%;  Score 48;  DB 1;  Length 205;
  Best Local Similarity   62.5%;  Pred. No. 5.6;
  Matches   10;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy         11 RELRRMSDEFEGSFKG 26
              ||:|||: | ||  ||
Db        168 REIRRMNKEQEGRSKG 183


RESULT 14
6PGL_THEMA
ID   6PGL_THEMA     STANDARD;      PRT;   220 AA.
AC   Q9X0N8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN   PGL OR DEVB OR TM1154.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC       PHOSPHOGLUCONATE.
CC   -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC       phospho-D-gluconate.
CC   -!- PATHWAY: Pentose phosphate pathway; second step.
CC   -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC       ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001772; AAD36230.1; -.
DR   PIR; F72289; F72289.
DR   TIGR; TM1154; -.
DR   InterPro; IPR006148; Gluc_gal_isom.
DR   InterPro; IPR005900; Phosphogluconlac.
DR   Pfam; PF01182; Glucosamine_iso; 1.
DR   TIGRFAMs; TIGR01198; pgl; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   220 AA;  25325 MW;  9B0FD07EE01E60C3 CRC64;

  Query Match             33.6%;  Score 48;  DB 1;  Length 220;
  Best Local Similarity   34.8%;  Pred. No. 6;
```

```
  Matches    8;  Conservative    8;  Mismatches    7;  Indels    0;  Gaps    0;

Qy          5 AAQRYGRELRRMSDEFEGSFKGL 27
              | ::| ||:|  :|:|: :  |:
Db        111 ACEKYEREIRSATDQFDLAILGM 133


RESULT 15
THRC_SOLTU
ID   THRC_SOLTU     STANDARD;      PRT;   519 AA.
AC   Q9MT28;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Threonine synthase, chloroplast precursor (EC 4.2.3.1) (TS).
OS   Solanum tuberosum (Potato).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; lamiids; Solanales; Solanaceae; Solanum.
OX   NCBI_TaxID=4113;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Leaf;
RX   PubMed=10940468;
RA   Casazza A.P., Basner A., Hoefgen R., Hesse H.;
RT   "Expression of threonine synthase from Solanum tuberosum L. is not
RT   metabolically regulated by photosynthesis-related signals or by
RT   nitrogenous compounds.";
RL   Plant Sci. 157:43-50(2000).
CC   -!- CATALYTIC ACTIVITY: O-phospho-L-homoserine + H(2)O = L-threonine +
CC       phosphate.
CC   -!- COFACTOR: Pyridoxal phosphate (By similarity).
CC   -!- ENZYME REGULATION: Allosterically activated by S-adenosyl-
CC       methionine (SAM) (By similarity).
CC   -!- PATHWAY: Threonine biosynthesis; last step.
CC   -!- SUBUNIT: Homodimer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Chloroplast (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE SERINE/THREONINE DEHYDRATASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF082894; AAF74984.1; -.
DR   HSSP; Q9S7B5; 1E5X.
DR   InterPro; IPR001926; B6_enzyme_beta.
DR   InterPro; IPR000634; S/T_dehydrtse.
DR   InterPro; IPR004450; Thr_synthase.
DR   Pfam; PF00291; PALP; 1.
DR   TIGRFAMs; TIGR00260; thrC; 1.
DR   PROSITE; PS00165; DEHYDRATASE_SER_THR; 1.
KW   Threonine biosynthesis; Lyase; Pyridoxal phosphate; Allosteric enzyme;
KW   Chloroplast; Transit peptide.
FT   TRANSIT       1     40       CHLOROPLAST (BY SIMILARITY).
FT   CHAIN        41    519       THREONINE SYNTHASE.
FT   BINDING     196    196       PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ   SEQUENCE   519 AA;  57412 MW;  114C0979CD231464 CRC64;

  Query Match             33.6%;  Score 48;  DB 1;  Length 519;
  Best Local Similarity   35.3%;  Pred. No. 15;
  Matches   12;  Conservative    6;  Mismatches    8;  Indels    8;  Gaps    1;

Qy          2 NLWAAQRYGRELRRMSD--------EFEGSFKGL 27
              ||: |:|:|::   |:|          |||| |
Db        165 NLFWAERFGKQFLGMTDLWVKHCGISHTGSFKDL 198


Search completed: September 15, 2003, 17:23:01
Job time : 6.36429 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 29.3143 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-56
Perfect score: 143
Sequence: 1 KNLWAAQRYGRELRRMSDEFEGSFKGL 27

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	87	60.8	146	13	Q9I9N2	Q9i9n2 brachydanio
2	53	37.1	196	16	Q8VJS3	Q8vjs3 mycobacteri
3	53	37.1	223	16	Q10843	Q10843 mycobacteri
4	53	37.1	946	11	Q8K016	Q8k016 mus musculu
5	52.5	36.7	505	8	O47148	O47148 menziesia c
6	52.5	36.7	506	8	O47149	O47149 rhododendro
7	52.5	36.7	506	8	O47171	O47171 rhododendro
8	52.5	36.7	506	8	O63960	O63960 rhododendro
9	52.5	36.7	506	8	O62982	O62982 rhododendro
10	52.5	36.7	506	8	O62975	O62975 rhododendro
11	52.5	36.7	506	8	O62972	O62972 rhododendro
12	52.5	36.7	506	8	O62989	O62989 rhododendro
13	52.5	36.7	506	8	O62978	O62978 rhododendro
14	52.5	36.7	506	8	O47155	O47155 rhododendro
15	52.5	36.7	506	8	O47152	O47152 rhododendro
16	52.5	36.7	506	8	O47173	O47173 rhododendro
17	52.5	36.7	506	8	O62990	O62990 rhododendro
18	52.5	36.7	506	8	O62974	O62974 rhododendro
19	52.5	36.7	506	8	O62993	O62993 menziesia m
20	52.5	36.7	506	8	O47170	O47170 rhododendro
21	52.5	36.7	506	8	O47174	O47174 rhododendro
22	52.5	36.7	506	8	O62983	O62983 rhododendro
23	52.5	36.7	506	8	O62980	O62980 rhododendro
24	52.5	36.7	506	8	O62981	O62981 rhododendro
25	52.5	36.7	506	8	O62977	O62977 rhododendro
26	52.5	36.7	506	8	O47168	O47168 menziesia p
27	52.5	36.7	506	8	O62988	O62988 rhododendro
28	52.5	36.7	506	8	O62973	O62973 rhododendro
29	52.5	36.7	506	8	O62992	O62992 ledum palus
30	52.5	36.7	506	8	O47175	O47175 rhododendro
31	52.5	36.7	506	8	Q8HSP1	Q8hsp1 rhododendro
32	52.5	36.7	506	8	Q8HSP0	Q8hsp0 rhododendro
33	52.5	36.7	506	8	Q8HSN9	Q8hsn9 rhododendro
34	52.5	36.7	506	8	Q8HSN8	Q8hsn8 rhododendro
35	52.5	36.7	506	8	Q8HSN7	Q8hsn7 rhododendro
36	52.5	36.7	506	8	Q8HSN6	Q8hsn6 rhododendro
37	52.5	36.7	506	8	Q8HSN5	Q8hsn5 rhododendro
38	52.5	36.7	506	8	Q8HSN4	Q8hsn4 rhododendro
39	52.5	36.7	507	8	O62985	O62985 rhododendro
40	52.5	36.7	507	8	O62986	O62986 rhododendro
41	52.5	36.7	508	8	O62979	O62979 rhododendro
42	52	36.4	471	17	Q8ZY71	Q8zy71 pyrobaculum
43	51.5	36.0	506	8	O47153	O47153 rhododendro
44	51.5	36.0	506	8	O47160	O47160 rhododendro
45	51.5	36.0	506	8	Q8HSP4	Q8hsp4 rhododendro

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 60.8%; Score 87; DB 13; Length 146;
Best Local Similarity 65.2%; Pred. No. 5.9e-05;
Matches 15; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 LWAAQRYGRELRRMSDEFEGSFK 25
              ||||::||::|||||||||:   |
Db         89 LWAAKKYGQQLRRMSDEFDKGMK 111
```

RESULT 2
Q8VJS3
ID Q8VJS3 PRELIMINARY; PRT; 196 AA.
AC Q8VJS3;

```
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   IS1607, transposase.
GN   MT2070.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CDC 1551 / Oshkosh;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA   Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA   Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA   Bishai W.;
RT   "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AE007058; AAK46348.1; -.
DR   TIGR; MT2070; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
SQ   SEQUENCE   196 AA;  21349 MW;  C145A8D836FD9C2D CRC64;

  Query Match             37.1%;  Score 53;  DB 16;  Length 196;
  Best Local Similarity   58.8%;  Pred. No. 8.4;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              ||||| || | : |  |
Db        134 NLWAADRYNRAIARGHD 150


RESULT 3
Q10843
ID   Q10843      PRELIMINARY;      PRT;   223 AA.
AC   Q10843;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Hypothetical protein Rv2014.
GN   RV2014 OR MTCY39.03C.
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX   NCBI_TaxID=1773;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=H37Rv;
RX   MEDLINE=98295987; PubMed=9634230;
RA   Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA   Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA   Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA   Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA   Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA   Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA   Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA   Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT   "Deciphering the biology of Mycobacterium tuberculosis from the
RT   complete genome sequence.";
RL   Nature 393:537-544(1998).
CC   -!- SIMILARITY: TO M.PARATUBERCULOSIS IS900.
DR   EMBL; Z74025; CAA98415.1; -.
DR   TubercuList; Rv2014; -.
DR   InterPro; IPR003346; Transposase_20.
DR   Pfam; PF02371; Transposase_20; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   223 AA;  24132 MW;  70456750017FEF37 CRC64;

  Query Match             37.1%;  Score 53;  DB 16;  Length 223;
  Best Local Similarity   58.8%;  Pred. No. 9.7;
  Matches   10;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSD 18
              ||||| || | : |  |
Db        165 NLWAADRYNRAIARGHD 181


RESULT 4
Q8K016
ID   Q8K016      PRELIMINARY;      PRT;   946 AA.
AC   Q8K016;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Inter-alpha trypsin inhibitor, heavy chain 2.
GN   ITIH2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Liver;
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC034341; AAH34341.1; -.
DR   MGD; MGI:96619; Itih2.
DR   InterPro; IPR006587; VIT.
DR   InterPro; IPR002035; VWF_A.
DR   SMART; SM00609; VIT; 1.
DR   SMART; SM00327; VWA; 1.
DR   PROSITE; PS50234; VWFA; 1.
SQ   SEQUENCE   946 AA;  105945 MW;  8B17DBA71B85BC5C CRC64;

  Query Match             37.1%;  Score 53;  DB 11;  Length 946;
  Best Local Similarity   34.6%;  Pred. No. 48;
  Matches    9;  Conservative    5;  Mismatches   12;  Indels    0;  Gaps    0;

Qy          2 NLWAAQRYGRELRRMSDEFEGSFKGL 27
              |:|  :  |     : | ||| |:|:
Db        212 NVWIMEPQGMRFLHVPDTFEGHFQGV 237


RESULT 5
O47148
ID   O47148      PRELIMINARY;      PRT;   505 AA.
AC   O47148;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K) (Fragment).
GN   MATK.
OS   Menziesia ciliicalyx.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Menziesia.
OX   NCBI_TaxID=49154;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61331; AAC15245.2; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
```

```
KW   mRNA processing; Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   505 AA;  60233 MW;  EE5F927AD2E57DE5 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 505;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        390 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 421


RESULT 6
O47149
ID   O47149      PRELIMINARY;      PRT;   506 AA.
AC   O47149;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron kiusianum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49167;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61332; AAB93753.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60502 MW;  0009EA88CD28549F CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 7
O47171
ID   O47171      PRELIMINARY;      PRT;   506 AA.
AC   O47171;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron edgeworthii.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49162;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61354; AAB93748.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60485 MW;  8A6353BFC5F4DC85 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 8
O63960
ID   O63960      PRELIMINARY;      PRT;   506 AA.
AC   O63960;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK OR YCF14.
OS   Rhododendron tashiroi, and
OS   Rhododendron farrerae.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75582, 75580;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012749; BAA25870.1; -.
DR   EMBL; AB012745; BAA25866.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60389 MW;  DE0C07AEE608B787 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 9
O62982
ID   O62982      PRELIMINARY;      PRT;   506 AA.
AC   O62982;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
```

```
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron nipponicum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75577;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012739; BAA25860.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60419 MW;  1F95132CCF4F6B40 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 10
O62975
ID   O62975      PRELIMINARY;      PRT;   506 AA.
AC   O62975;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron ponticum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49628;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012732; BAA25853.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60449 MW;  21DFF700B071B5B8 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 11
O62972
ID   O62972      PRELIMINARY;      PRT;   506 AA.
AC   O62972;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron ovatum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49169;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012729; BAA25850.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60493 MW;  D230E54B8C20FEF0 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 12
O62989
ID   O62989      PRELIMINARY;      PRT;   506 AA.
AC   O62989;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron indicum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=75581;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
```

```
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012747; BAA25868.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60489 MW;  6D38A1D4D6FEC9BF CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 13
O62978
ID   O62978      PRELIMINARY;      PRT;   506 AA.
AC   O62978;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron canadense.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49465;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kurashige Y., Mine M., Kobayashi N., Handa T., Takayanagi K.,
RA   Yukawa T.;
RT   "Investigation of Sectional Relationships in the Genus
RT   Rhododendron(Ericaceae) based on matK Sequences.";
RL   J. Jpn. Bot. 0:0-0(1998).
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; AB012735; BAA25856.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60350 MW;  5E832589ED64EA25 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 14
O47155
ID   O47155      PRELIMINARY;      PRT;   506 AA.
AC   O47155;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron hongkongense.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49165;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61338; AAB93751.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60406 MW;  4B5C675CE32218D8 CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
  Matches   12;  Conservative    5;  Mismatches    8;  Indels    7;  Gaps    1;

Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : || |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422


RESULT 15
O47152
ID   O47152      PRELIMINARY;      PRT;   506 AA.
AC   O47152;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Ribosomal maturase (Intron maturase) (Maturase K).
GN   MATK.
OS   Rhododendron tomentosum.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Asteridae; Ericales; Ericaceae; Ericoideae; Rhodoreae; Rhododendron.
OX   NCBI_TaxID=49170;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kron K.A.;
RT   "Phylogenetics of Rhododendroideae (Ericaceae).";
RL   Submitted (JUN-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PROBABLY ASSISTS IN SPLICING CHLOROPLAST GROUP II
CC       INTRONS (BY SIMILARITY).
CC   -!- SIMILARITY: WITH CORRESPONDING ORF IN OTHER PLANT CHLOROPLASTS,
CC       AND REGIONS OF SIMILARITY TO MATURASE-LIKE POLYPEPTIDES ENCODED BY
CC       MITOCHONDRIAL INTRONS.
DR   EMBL; U61335; AAB93757.1; -.
DR   InterPro; IPR000442; Intron_maturse2.
DR   InterPro; IPR002866; MatK_N.
DR   Pfam; PF01348; Intron_maturas2; 1.
DR   Pfam; PF01824; MatK_N; 1.
KW   mRNA processing; Chloroplast.
SQ   SEQUENCE   506 AA;  60439 MW;  855FDBDB8A5F800D CRC64;

  Query Match             36.7%;  Score 52.5;  DB 8;  Length 506;
  Best Local Similarity   37.5%;  Pred. No. 28;
```

Matches 12; Conservative 5; Mismatches 8; Indels 7; Gaps 1;

```
Qy          1 KNLWAA-------QRYGRELRRMSDEFEGSFK 25
              | :|||       :|:||  | :|  : ||| |
Db        391 KPVWAALSDSDIIERFGRIYRNLSHYYSGSLK 422
                                    |
```

Search completed: September 15, 2003, 17:25:50
Job time : 29.3143 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 22.6286 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	100.0	16	20	AAY05422	Mouse BAD BH3 doma
2	84	100.0	16	21	AAB37028	Bcl2 polypeptide B
3	84	100.0	23	17	AAR95166	bcl-x(L)/bcl-2 ass
4	84	100.0	26	21	AAB37001	Bcl2 polypeptide B
5	84	100.0	26	21	AAB37002	Bcl2 polypeptide B
6	84	100.0	27	21	AAB37003	Bcl2 polypeptide B
7	84	100.0	27	21	AAB37056	Bcl2 polypeptide B
8	84	100.0	28	21	AAB37055	Bcl2 polypeptide B
9	84	100.0	162	22	AAB70370	Shorter murine BAD
10	84	100.0	204	17	AAR95168	bcl-x(L)/bcl-2 ass
11	84	100.0	204	19	AAW61315	Murine BCL-XL/BCL-
12	84	100.0	204	19	AAW61316	Mutant BCL-XL/BCL-
13	84	100.0	204	19	AAW61317	Mutant BCL-XL/BCL-
14	84	100.0	204	19	AAW61318	Mutant BCL-XL/BCL-
15	84	100.0	204	19	AAW58832	Murine BAD protein
16	84	100.0	204	22	AAB70369	Longer murine BAD
17	84	100.0	204	24	ABR39082	Murine BAD protein
18	84	100.0	567	22	AAU00220	Bad-DTTR apoptosis
19	73	86.9	16	20	AAY05421	Human BAD BH3 doma
20	73	86.9	16	21	AAB37029	Bcl2 polypeptide B
21	73	86.9	18	23	ABG78483	Human Bcl2 fluores
22	73	86.9	18	23	AAU78609	Human Bad peptide
23	73	86.9	20	23	ABG78499	Mutant Bcl2 compet
24	73	86.9	20	23	AAU78626	Human Bad peptide
25	73	86.9	21	23	ABG78500	Mutant Bcl2 compet
26	73	86.9	21	23	AAU78608	Human Bad peptide
27	73	86.9	21	23	AAU78630	Human Bad peptide
28	73	86.9	22	23	AAU78629	Human Bad peptide
29	73	86.9	23	23	AAU78628	Human Bad peptide
30	73	86.9	24	23	ABG78482	Human Bcl2 fluores
31	73	86.9	24	23	AAU78605	Human Bad peptide
32	73	86.9	24	23	AAU78627	Human Bad peptide
33	73	86.9	25	23	ABP56161	PTPC-interacting T
34	73	86.9	25	23	ABG78481	Human Bcl2 fluores
35	73	86.9	25	23	ABG78484	Mutant Bcl2 compet
36	73	86.9	25	23	ABG78485	Mutant Bcl2 compet
37	73	86.9	25	23	ABG78486	Mutant Bcl2 compet
38	73	86.9	25	23	ABG78487	Mutant Bcl2 compet
39	73	86.9	25	23	ABG78488	Mutant Bcl2 compet
40	73	86.9	25	23	ABG78489	Mutant Bcl2 compet
41	73	86.9	25	23	ABG78490	Mutant Bcl2 compet
42	73	86.9	25	23	ABG78491	Mutant Bcl2 compet
43	73	86.9	25	23	ABG78492	Mutant Bcl2 compet
44	73	86.9	25	23	ABG78493	Mutant Bcl2 compet
45	73	86.9	25	23	ABG78494	Mutant Bcl2 compet

ALIGNMENTS

RESULT 1
AAY05422
ID AAY05422 standard; peptide; 16 AA.
XX
AC AAY05422;
XX
DT 02-JUL-1999 (first entry)
XX
DE Mouse BAD BH3 domain.
XX
KW BH3 domain; cell death agonist; bcl homology domain; BCL-2 family;
KW apoptosis promoter; cancer cell; virus infected cell; inflammation;
KW autoantibody producing cell; cancer; lymphoproliferative condition;
KW arthritis; autoimmune disease; therapy.
XX
OS Mus sp.
XX
PN WO9916787-A1.
XX
PD 08-APR-1999. 10/2/01 (1 day after App. PD)
XX
PF 22-SEP-1998; 98WO-US19765.
XX
PR 07-OCT-1997; 97US-0946039.
PR 26-SEP-1997; 97US-0060133.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1999-255058/21.
XX

PT Bcl homology domain 3 polypeptide
XX
PS Example 1; Fig 4; 104pp; English.
XX
CC This sequence represents the BH3 domain of mouse BAD.
CC The invention relates to a bcl homology domain 3 (BH3 domain),
CC derived from a proapoptotic member of the BCL-2 family. The
CC BH3 polypeptide can be used in a method for promoting apoptosis in a
CC target cell, especially where the cell is a cancer cell a virus infected
CC cell or an autoantibody producing cell. The BH3 polypeptide can be used
CC in therapeutic compositions for treating disease including cancer, other
CC lymphoproliferative conditions, arthritis, inflammation, and autoimmune
CC diseases, which may result from the down regulation of cell death
CC regulation.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 84; DB 20; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 1 QRYGRELRRMSDEFEG 16

RESULT 2
AAB37028
ID AAB37028 standard; peptide; 16 AA.
XX
AC AAB37028;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #28.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 18; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C_t-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 84; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 2.6e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 1 QRYGRELRRMSDEFEG 16

RESULT 3
AAR95166
ID AAR95166 standard; peptide; 23 AA.
XX
AC AAR95166;
XX
DT 03-JAN-1997 (first entry)
XX
DE bcl-x(L)/bcl-2 associated death promoter epitope, residues 138-160.
XX
KW Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS Mus musculus.
XX
PN WO9613614-A1.
XX
PD 09-MAY-1996.
XX
PF 31-OCT-1995; 95WO-US14246.
XX
PR 31-OCT-1994; 94US-0333565.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1996-251465/25.
XX
PT Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT useful to treat neoplasia and apoptosis and to identify agents
PT inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS Claim 2; Page 103; 130pp; English.
XX
CC The sequences given in AAR95155-67 represent epitopes derived from the
CC murine bcl-x(L)/bcl-2 associated death promoter (Bad) polypeptide (see
CC also AAR95168). Bad is a 22.1 kD protein which interacts with bcl-2 and
CC bcl-x proteins and regulates cell death. It has homology to the bcl-2-
CC related family clustered in the BH1 and BH2 domain. Bad has been found
CC to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid assays and in
CC vivo in mammalian cells. Overexpressed Bad counters the death

CC inhibitory activity of bcl-x(L), but is much less effective at countering
CC the death inhibitory activity of bcl-2. Bad expression can accelerate
CC apoptotic cell death induced by cytokine deprivation in an IL-3 dependent
CC cell line expressing bcl-x(L), and its also counters the death repressor
CC activity of bcl-x(L). Bad competes with Bax for binding to bcl-x(L).
CC Bad may be used to identify agents which inhibit its binding to bcl-2
CC or bcl-x(L) to form heterodimers. Such agents may be used to treat
CC neurodegenerative diseases, immunodeficiency diseases, e.g. AIDS,
CC senescence or ischaemia.
XX
SQ Sequence 23 AA;

Query Match 100.0%; Score 84; DB 17; Length 23;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 8 QRYGRELRRMSDEFEG 23

RESULT 4
AAB37001
ID AAB37001 standard; peptide; 26 AA.
XX
AC AAB37001;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #1.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 84; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFEG 21

RESULT 5
AAB37002
ID AAB37002 standard; peptide; 26 AA.
XX
AC AAB37002;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #2.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain

CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 26 AA;

Query Match 100.0%; Score 84; DB 21; Length 26;
Best Local Similarity 100.0%; Pred. No. 4.2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFEG 21

RESULT 6
AAB37003
ID AAB37003 standard; peptide; 27 AA.
XX
AC AAB37003;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #3.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 17; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 84; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFEG 21

RESULT 7
AAB37056
ID AAB37056 standard; peptide; 27 AA.
XX
AC AAB37056;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #56.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,

CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 27 AA;

Query Match 100.0%; Score 84; DB 21; Length 27;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 7 QRYGRELRRMSDEFEG 22

RESULT 8
AAB37055
ID AAB37055 standard; peptide; 28 AA.
XX
AC AAB37055;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #55.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 19; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 28 AA;

Query Match 100.0%; Score 84; DB 21; Length 28;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 7 QRYGRELRRMSDEFEG 22

RESULT 9
AAB70370
ID AAB70370 standard; protein; 162 AA.
XX
AC AAB70370;
XX
DT 02-MAY-2001 (first entry)
XX
DE Shorter murine BAD mutant amino acid sequence SEQ ID NO:3.
XX
KW Bcl-XL/Bcl-2 associated cell death regulator; BAD; mutant; apoptosis;
KW immunostimulant; neuroprotective; nootropic; antiischaemic; vulnerary;
KW cytostatic; antiviral; antiarthritic; antiinflammatory; wound healing;
KW immunosuppressive; apoptosis inducer; apoptosis inhibitor; cancer;
KW immunodeficiency disease; neurodegenerative disease; viral infection;
KW ischaemic cell death; reperfusion cell death; arthritis; infertility;
KW lymphoproliferative condition; inflammation; autoimmune disease.
XX
OS Mus musculus.
OS Synthetic.
XX
PN WO200110888-A1.
XX
PD 15-FEB-2001.
XX
PF 30-MAY-2000; 2000WO-US11864.
XX
PR 28-MAY-1999; 99US-0136783.
XX
PA (APOP-) APOPTOSIS TECHNOLOGY INC.
XX
PI Zhou X;
XX
DR WPI; 2001-138734/14.
XX
PT New mutant Bcl-XL/Bcl-2 Associated Cell Death Regulator polypeptide,
PT useful for screening for candidate compounds which induce or inhibit

PT apoptosis, comprises amino acid substitutions at Ser118, Ser155 or
PT Ser113 -
XX
PS Claim 7; Page 148-149; 157pp; English.
XX
CC The present invention describes an isolated or synthetic polypeptide
CC (I) comprising a less than full length amino acid sequence of a mutant
CC Bcl-XL/Bcl-2 associated cell death regulator polypeptide (BAD) or its
CC fragment, which contains amino acid substitutions at Ser118 of a human
CC BAD, Ser155 of a murine BAD (longer murine BAD) or Ser113 of a murine
CC BAD (shorter murine BAD). (I) has immunostimulant, neuroprotective,
CC nootropic, antiischaemic, vulnerary, cytostatic, antiviral,
CC antiarthritic, antiinflammatory and immunosuppressive activities, and
CC can be used as an apoptosis inducer or inhibitor. BAD polypeptides and
CC polynucleotides can be used for screening candidate compounds and drugs
CC for activity that promote cell survival or apoptosis. Other uses include
CC inducing or inhibiting apoptosis in a cell. Candidate compounds
CC identified and (mutant) BAD polypeptides are useful in treating
CC immunodeficiency diseases, neurodegenerative diseases, ischaemic cell
CC death, reperfusion cell death, wound healing, cancer, viral infections,
CC lymphoproliferative conditions, arthritis, infertility, inflammation and
CC autoimmune diseases. The present sequence represents a specifically
CC claimed shorter murine BAD mutant amino acid sequence from the present
CC invention.
XX
SQ Sequence 162 AA;

Query Match 100.0%; Score 84; DB 22; Length 162;
Best Local Similarity 100.0%; Pred. No. 2.7e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 103 QRYGRELRRMSDEFEG 118

RESULT 10
AAR95168
ID AAR95168 standard; Protein; 204 AA.
XX
AC AAR95168;
XX
DT 06-JAN-1997 (first entry)
XX
DE bcl-x(L)/bcl-2 associated death promoter protein.
XX
KW Epitope; murine; bcl-x(L)/bcl-2 associated death promoter; Bad; stroke;
KW polypeptide; bcl-x; cell death; regulate; BH1; BH2; apoptotic cell death;
KW cytokine deprivation; IL-3 dependent cell line; immunodeficiency; AIDS;
KW neurodegenerative disease; senescence; ischaemia; neoplasia.
XX
OS Mus musculus.
XX
FH Key Location/Qualifiers
FT Region 147..149
FT /note= "BH1 conserved amino acids"
FT Region 191..192
FT /note= "BH2 conserved amino acids"
FT Domain 38..61
FT /note= "PEST sequence"
FT Domain 111..130
FT /note= "PEST sequence"
XX
PN WO9613614-A1.
XX
PD 09-MAY-1996.
XX
PF 31-OCT-1995; 95WO-US14246.
XX
PR 31-OCT-1994; 94US-0333565.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1996-251465/25.
DR N-PSDB; AAT29479.
XX
PT Polynucleotide encoding bcl-x(L)/bcl-2 associated death promoter -
PT useful to treat neoplasia and apoptosis and to identify agents
PT inhibiting its binding to bcl-2 or bcl-x(L) to form heteromultimers
XX
PS Claim 3; Fig 1; 130pp; English.
XX
CC This sequence represents the murine bcl-x(L)/bcl-2 associated death
CC promoter (Bad) gene. Bad is a 22.1 kD protein which interacts with
CC bcl-2 and bcl-x proteins and regulates cell death. It has homology
CC to the bcl-2-related family clustered in the BH1 and BH2 domain. Bad
CC has been found to hybridise to bcl-x(L) and bcl-2 in yeast two-hybrid
CC assays and in vivo in mammalian cells. Overexpressed Bad counters the
CC death inhibitory activity of bcl-x(L), but is much less effective at
CC countering the death inhibitory activity of bcl-2. Bad expression can
CC accelerate apoptotic cell death induced by cytokine deprivation in an
CC IL-3 dependent cell line expressing bcl-x(L), and its also counters the
CC death repressor activity of bcl-x(L). Bad competes with Bax for binding
CC to bcl-x(L). Bad may be used to identify agents which inhibit its
CC binding to bcl-2 or bcl-x(L) to form heterodimers. Such agents may be
CC used to treat neurodegenerative diseases, immunodeficiency diseases,
CC e.g. AIDS, senescence or ischaemia.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 84; DB 17; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 145 QRYGRELRRMSDEFEG 160

RESULT 11
AAW61315
ID AAW61315 standard; Protein; 204 AA.
XX
AC AAW61315;
XX
DT 07-OCT-1998 (first entry)
XX
DE Murine BCL-XL/BCL-2 associated cell death regulator.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27833.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 1; Fig 10; 95pp; English.

XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence is the murine BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 84; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 145 QRYGRELRRMSDEFEG 160

RESULT 12
AAW61316
ID AAW61316 standard; Protein; 204 AA.
XX
AC AAW61316;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #1.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27834.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 59; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC death regulator) proteins, having an amino acid other than Ser at
CC position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC present sequence represents a mutant BAD protein. Also described are: (1)
CC fragments of mutant BAD protein able to decrease cell viability; (2)
CC fusion proteins of mutant BAD with a heterologous polypeptide that
CC increases intracellular delivery. Mutant BAD proteins are used to treat
CC or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC viral infection, lymphoproliferation, arthritis, infertility,
CC inflammation and autoimmune disease. Polynucleotide sequences encoding
CC mutant BAD proteins can be used similarly by gene therapy or to produce
CC transgenic animals for use as disease models or in drug screening. BAD
CC proteins phosphorylated at specified Ser are used to screen for enhancers
CC and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC aging or ischaemic cell death. The apoptotic status of cells is
CC determined by measuring relative amounts of phosphorylated and non-
CC phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC greater death-promoting activity than wild-type BAD which can become
CC phosphorylated on the specified Ser, forming a product that does not
CC heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC proteins in the cytosol, thus promoting cell survival. The mutants with
CC Ser substituted cannot bind 14-3-3.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 84; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 145 QRYGRELRRMSDEFEG 160

RESULT 13
AAW61317
ID AAW61317 standard; Protein; 204 AA.
XX
AC AAW61317;
XX
DT 07-OCT-1998 (first entry)
XX
DE Mutant BCL-XL/BCL-2 associated cell death regulator #2.
XX
KW Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS Mus sp.
OS Synthetic.
XX
PN WO9817682-A1.
XX
PD 30-APR-1998.
XX
PF 17-OCT-1997; 97WO-US19175.
XX
PR 18-OCT-1996; 96US-0733505.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1998-261422/23.
DR N-PSDB; AAV27835.
XX
PT New mutant BAD polypeptide with phosphorylatable serine replaced -
PT useful for, e.g. treating reduced apoptosis such as in cancer or
PT viral infection
XX
PS Claim 7; Page 60; 95pp; English.
XX
CC The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell

```
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 84;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 14
AAW61318
ID   AAW61318 standard; Protein; 204 AA.
XX
AC   AAW61318;
XX
DT   07-OCT-1998  (first entry)
XX
DE   Mutant BCL-XL/BCL-2 associated cell death regulator #3.
XX
KW   Murine; mouse; BCL-XL/BCL-2 associated cell death regulator; BAD protein;
KW   serine substituted mutant; apoptosis; cancer; viral infection.
XX
OS   Mus sp.
OS   Synthetic.
XX
PN   WO9817682-A1.
XX
PD   30-APR-1998.
XX
PF   17-OCT-1997;   97WO-US19175.
XX
PR   18-OCT-1996;   96US-0733505.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-261422/23.
DR   N-PSDB; AAV27836.
XX
PT   New mutant BAD polypeptide with phosphorylatable serine replaced -
PT   useful for, e.g. treating reduced apoptosis such as in cancer or
PT   viral infection
XX
PS   Claim 7; Page 60-61; 95pp; English.
XX
CC   The present invention describes mutant BAD (BCL-XL/BCL-2 associated cell
CC   death regulator) proteins, having an amino acid other than Ser at
CC   position 112 and/or 136, relative to the murine BAD 204 aa sequence. The
CC   present sequence represents a mutant BAD protein. Also described are: (1)
CC   fragments of mutant BAD protein able to decrease cell viability; (2)
CC   fusion proteins of mutant BAD with a heterologous polypeptide that
CC   increases intracellular delivery. Mutant BAD proteins are used to treat
CC   or prevent diseases associated with reduced apoptosis, e.g. cancer,
CC   viral infection, lymphoproliferation, arthritis, infertility,
CC   inflammation and autoimmune disease. Polynucleotide sequences encoding
CC   mutant BAD proteins can be used similarly by gene therapy or to produce
CC   transgenic animals for use as disease models or in drug screening. BAD
CC   proteins phosphorylated at specified Ser are used to screen for enhancers
CC   and inhibitors of serine-phosphatase. Inhibitors are potentially useful
CC   in treatment of excessive apoptosis such as AIDS, neurodegeneration,
CC   aging or ischaemic cell death. The apoptotic status of cells is
CC   determined by measuring relative amounts of phosphorylated and non-
CC   phosphorylated BAD, by usual immunoassays. Mutant BAD proteins have
CC   greater death-promoting activity than wild-type BAD which can become
CC   phosphorylated on the specified Ser, forming a product that does not
CC   heterodimerise with BCL-2 or BCL-XL but instead binds to 14-3-3 family
CC   proteins in the cytosol, thus promoting cell survival. The mutants with
CC   Ser substituted cannot bind 14-3-3.
XX
SQ   Sequence   204 AA;

  Query Match             100.0%;  Score 84;  DB 19;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 3.5e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 15
AAW58832
ID   AAW58832 standard; protein; 204 AA.
XX
AC   AAW58832;
XX
DT   23-JUL-1998  (first entry)
XX
DE   Murine BAD protein.
XX
KW   BAD protein; Bcl-XL/Bcl-2 associated cell death regulator; 14-3-3;
KW   serine phosphorylation; post-translational modification; apoptosis;
KW   signal transduction regulator; phosphoserine phosphatase; senescence;
KW   immunodeficiency disease; neurodegenerative disease; infertility;
KW   cancer, viral infection; lymphoproliferative condition; arthritis;
KW   inflammation; autoimmune diseases.
XX
OS   Mus sp.
XX
PN   WO9809643-A1.
XX
PD   12-MAR-1998.
XX
PF   09-SEP-1997;   97WO-US15871.
XX
PR   09-SEP-1996;   96US-0707868.
XX
PA   (UNIW ) UNIV WASHINGTON.
XX
PI   Korsmeyer SJ;
XX
DR   WPI; 1998-207049/18.
XX
PT   Serine-phosphorylated Bcl-X-l/Bcl-2 Associated cell Death regulator
PT   polypeptide - useful for modulation of apoptosis associated with,
PT   e.g. cancer and immunodeficiency diseases
XX
PS   Claim 3; Fig 8; 61pp; English.
XX
CC   This sequence represents a novel serine-phosphorylated protein, BAD
```

CC (Bcl-Xl/Bcl-2 associated cell death regulator). The serine residue is
CC phosporylated in a post-translational modification and allows binding
CC to the 14-3-3 protein which is a signal transduction regulator.
CC Modulators of phosphorylated BAD, which act through inhibition/activation
CC of a phosphoserine phosphatase, are useful for preventing/treating
CC increased/decreased apoptosis in a cell. The increased apoptosis may
CC result from immunodeficiency diseases, senescence, neurodegenerative
CC disease, ischaemic cell death, reperfusion cell death, infertility and
CC wound-healing. Decreased apoptosis may result from cancer, viral
CC infection, lymphoproliferative conditions, arthritis, infertility,
CC inflammation and autoimmune diseases. Measuring the amount of
CC phosphorylated compared to unphosphorylated BAD polypeptide and/or total
CC BAD in a cell is useful for determining the apoptotic state of a cell.
XX
SQ Sequence 204 AA;

Query Match 100.0%; Score 84; DB 19; Length 204;
Best Local Similarity 100.0%; Pred. No. 3.5e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160
```

Search completed: September 15, 2003, 17:22:13
Job time : 22.6286 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 8.34286 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	100.0	23	1	US-08-333-565-10	Sequence 10, Appl
2	84	100.0	23	2	US-08-661-479-10	Sequence 10, Appl
3	84	100.0	204	1	US-08-333-565-2	Sequence 2, Appli
4	84	100.0	204	2	US-08-661-479-2	Sequence 2, Appli
5	84	100.0	204	2	US-08-733-505A-1	Sequence 1, Appli
6	84	100.0	204	2	US-08-733-505A-12	Sequence 12, Appl
7	84	100.0	204	2	US-08-733-505A-13	Sequence 13, Appl
8	84	100.0	204	2	US-08-733-505A-14	Sequence 14, Appl
9	81	96.4	204	2	US-08-717-123-3	Sequence 3, Appli
10	81	96.4	204	4	US-09-375-257-3	Sequence 3, Appli
11	73	86.9	59	2	US-08-733-505A-55	Sequence 55, Appl
12	73	86.9	59	2	US-08-733-505A-56	Sequence 56, Appl
13	73	86.9	59	2	US-08-733-505A-57	Sequence 57, Appl
14	73	86.9	59	2	US-08-733-505A-58	Sequence 58, Appl
15	73	86.9	166	1	US-08-665-617-2	Sequence 2, Appli
16	73	86.9	168	2	US-08-717-123-2	Sequence 2, Appli
17	73	86.9	168	3	US-08-985-335-1	Sequence 1, Appli
18	73	86.9	168	3	US-08-985-335-7	Sequence 7, Appli
19	73	86.9	168	3	US-09-410-372-1	Sequence 1, Appli
20	73	86.9	168	3	US-09-410-372-7	Sequence 7, Appli
21	73	86.9	168	4	US-09-375-257-2	Sequence 2, Appli
22	67	79.8	16	1	US-08-333-565-26	Sequence 26, Appl
23	67	79.8	16	2	US-08-661-479-26	Sequence 26, Appl
24	43	51.2	213	3	US-08-718-738-18	Sequence 18, Appl
25	43	51.2	213	3	US-09-221-844-18	Sequence 18, Appl
26	43	51.2	591	4	US-09-252-991A-27090	Sequence 27090, A
27	42	50.0	11	2	US-08-733-505A-34	Sequence 34, Appl
28	42	50.0	11	2	US-08-706-741B-69	Sequence 69, Appl
29	42	50.0	11	2	US-08-924-695A-69	Sequence 69, Appl
30	41	48.8	322	4	US-09-359-161-7	Sequence 7, Appli
31	40.5	48.2	904	4	US-09-328-352-4656	Sequence 4656, Ap
32	40	47.6	505	4	US-09-252-991A-21995	Sequence 21995, A
33	39	46.4	66	2	US-08-867-087B-40	Sequence 40, Appl
34	39	46.4	333	4	US-09-252-991A-28443	Sequence 28443, A
35	39	46.4	1125	4	US-09-252-991A-18729	Sequence 18729, A
36	38.5	45.8	1112	4	US-09-252-991A-27256	Sequence 27256, A
37	38	45.2	376	4	US-09-252-991A-26270	Sequence 26270, A
38	38	45.2	574	4	US-09-252-991A-29351	Sequence 29351, A
39	38	45.2	670	4	US-09-252-991A-26867	Sequence 26867, A
40	37.5	44.6	906	4	US-09-252-991A-31458	Sequence 31458, A
41	37	44.0	125	4	US-09-328-352-7449	Sequence 7449, Ap
42	37	44.0	143	4	US-09-252-991A-29777	Sequence 29777, A
43	37	44.0	284	4	US-09-328-352-6559	Sequence 6559, Ap
44	37	44.0	301	4	US-09-252-991A-26483	Sequence 26483, A
45	37	44.0	309	4	US-09-252-991A-32234	Sequence 32234, A

ALIGNMENTS

```
RESULT 1
US-08-333-565-10
; Sequence 10, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-333-565-10

  Query Match            100.0%;  Score 84;  DB 1;  Length 23;
  Best Local Similarity  100.0%;  Pred. No. 5.4e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db          8 QRYGRELRRMSDEFEG 23
```

```
RESULT 2
US-08-661-479-10
; Sequence 10, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-661-479-10

  Query Match             100.0%;  Score 84;  DB 2;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 5.4e-08;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db          8 QRYGRELRRMSDEFEG 23


RESULT 3
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-333-565-2

  Query Match             100.0%;  Score 84;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 4
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
```

```
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             100.0%;  Score 84;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 5
US-08-733-505A-1
; Sequence 1, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-733-505A-1

  Query Match             100.0%;  Score 84;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 6
US-08-733-505A-12
; Sequence 12, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  12:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-12

  Query Match             100.0%;  Score 84;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 7
US-08-733-505A-13
; Sequence 13, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
```

```
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  13:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-13

  Query Match            100.0%;  Score 84;  DB 2;  Length 204;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 8
US-08-733-505A-14
; Sequence 14, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  14:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-14

  Query Match            100.0%;  Score 84;  DB 2;  Length 204;
  Best Local Similarity  100.0%;  Pred. No. 5.5e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 9
US-08-717-123-3
; Sequence 3, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  3:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
US-08-717-123-3

  Query Match            96.4%;  Score 81;  DB 2;  Length 204;
  Best Local Similarity  93.8%;  Pred. No. 1.8e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              |||||||||| :|||||
Db        145 QRYGRELRRMTDEFEG 160


RESULT 10
US-09-375-257-3
; Sequence 3, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-375-257-3

  Query Match            96.4%;  Score 81;  DB 4;  Length 204;
  Best Local Similarity  93.8%;  Pred. No. 1.8e-06;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||:|||||
Db        145 QRYGRELRRMTDEFEG 160


RESULT 11
US-08-733-505A-55
; Sequence 55, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-55

  Query Match             86.9%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 12
US-08-733-505A-56
; Sequence 56, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-56

  Query Match             86.9%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 13
US-08-733-505A-57
; Sequence 57, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-57
```

```
  Query Match           86.9%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 14
US-08-733-505A-58
; Sequence 58, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-58

  Query Match           86.9%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity  100.0%;  Pred. No. 1.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 15
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match           86.9%;  Score 73;  DB 1;  Length 166;
  Best Local Similarity  100.0%;  Pred. No. 3.4e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        106 QRYGRELRRMSDEF 119


Search completed:  September 15, 2003, 17:45:06
Job time :  9.34286 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 12.6857 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	81	96.4	204	9	US-09-922-378-3	Sequence 3, Appli
2	81	96.4	204	14	US-10-066-179-3	Sequence 3, Appli
3	73	86.9	25	15	US-10-059-261-258	Sequence 258, App
4	73	86.9	168	9	US-09-922-378-2	Sequence 2, Appli
5	73	86.9	168	9	US-09-894-657-1	Sequence 1, Appli
6	73	86.9	168	9	US-09-894-657-7	Sequence 7, Appli
7	73	86.9	168	14	US-10-066-179-2	Sequence 2, Appli
8	67	79.8	15	15	US-10-174-105A-147	Sequence 147, App
9	50	59.5	215	15	US-10-156-761-9145	Sequence 9145, Ap
10	43	51.2	213	9	US-09-843-846-18	Sequence 18, Appl
11	41	48.8	582	10	US-09-331-631A-22	Sequence 22, Appl
12	40	47.6	380	15	US-10-156-761-13594	Sequence 13594, A
13	40	47.6	2871	15	US-10-146-473-41	Sequence 41, Appl
14	38.5	45.8	764	15	US-10-166-087-16	Sequence 16, Appl
15	38	45.2	35	15	US-10-092-750-1	Sequence 1, Appli
16	38	45.2	50	10	US-09-971-980-64	Sequence 64, Appl
17	38	45.2	64	10	US-09-971-980-62	Sequence 62, Appl
18	38	45.2	138	15	US-10-092-750-241	Sequence 241, App
19	38	45.2	327	15	US-10-102-806-570	Sequence 570, App
20	38	45.2	385	15	US-10-146-772-122	Sequence 122, App
21	38	45.2	571	9	US-09-815-242-11813	Sequence 11813, A
22	38	45.2	1053	9	US-09-841-132-583	Sequence 583, App
23	38	45.2	1053	14	US-10-007-693-97	Sequence 97, Appl
24	38	45.2	4840	15	US-10-156-761-10435	Sequence 10435, A
25	37.5	44.6	1265	15	US-10-198-070-69	Sequence 69, Appl
26	37	44.0	146	15	US-10-121-757B-12	Sequence 12, Appl
27	37	44.0	198	10	US-09-738-626-4812	Sequence 4812, Ap
28	37	44.0	334	10	US-09-794-715A-8	Sequence 8, Appli
29	37	44.0	334	12	US-10-286-581-8	Sequence 8, Appli
30	37	44.0	334	15	US-10-046-924-8	Sequence 8, Appli
31	37	44.0	426	9	US-09-815-242-5704	Sequence 5704, Ap
32	37	44.0	485	14	US-10-047-412A-4	Sequence 4, Appli
33	37	44.0	495	15	US-10-156-761-10045	Sequence 10045, A
34	37	44.0	515	15	US-10-106-698-4658	Sequence 4658, Ap
35	37	44.0	556	12	US-09-949-029-36	Sequence 36, Appl
36	37	44.0	575	9	US-09-839-136-2	Sequence 2, Appli
37	37	44.0	575	9	US-09-839-136-10	Sequence 10, Appl
38	37	44.0	705	9	US-09-815-242-12463	Sequence 12463, A
39	36.5	43.5	817	10	US-09-992-481-4	Sequence 4, Appli
40	36	42.9	61	9	US-09-925-301-1577	Sequence 1577, Ap
41	36	42.9	78	15	US-10-083-357-1237	Sequence 1237, Ap
42	36	42.9	80	15	US-10-106-698-5539	Sequence 5539, Ap
43	36	42.9	87	9	US-09-925-301-1433	Sequence 1433, Ap
44	36	42.9	95	10	US-09-925-300-1788	Sequence 1788, Ap
45	36	42.9	113	10	US-09-925-300-1789	Sequence 1789, Ap

ALIGNMENTS

```
RESULT 1
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-922-378-3

  Query Match            96.4%;  Score 81;  DB 9;  Length 204;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-05;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||:|||||
Db        145 QRYGRELRRMTDEFEG 160


RESULT 2
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
```

```
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match             96.4%;  Score 81;  DB 14;  Length 204;
  Best Local Similarity   93.8%;  Pred. No. 1.4e-05;
  Matches   15;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||:|||||
Db        145 QRYGRELRRMTDEFEG 160


RESULT 3
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
;  APPLICANT: EDELMAN, LENA
;  APPLICANT:  JACOTOT, ETIENNE DANIEL FRANCOIS
;  APPLICANT:  BRIAND, JEAN-PAUL
;  TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
;  TITLE OF INVENTION:  SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
;  TITLE OF INVENTION:  FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
;  TITLE OF INVENTION:  (PTPC)
;  FILE REFERENCE: 03495.0216
;  CURRENT APPLICATION NUMBER: US/10/059,261
;  CURRENT FILING DATE:  2002-08-29
;  PRIOR APPLICATION NUMBER: 60/265,594
;  PRIOR FILING DATE: 2001-02-02
;  NUMBER OF SEQ ID NOS: 325
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
;   LENGTH: 25
;   TYPE: PRT
;   ORGANISM: Unknown Organism
;   FEATURE:
;   OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

  Query Match             86.9%;  Score 73;  DB 15;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 3.1e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db          6 QRYGRELRRMSDEF 19


RESULT 4
US-09-922-378-2
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match             86.9%;  Score 73;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        108 QRYGRELRRMSDEF 121


RESULT 5
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: SYNORAB01
;            CLONE: 358673
;       SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match             86.9%;  Score 73;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        108 QRYGRELRRMSDEF 121


RESULT 6
US-09-894-657-7
```

```
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;  GENERAL INFORMATION:
;       APPLICANT: Hillman, Jennifer L.
;                  Yue, Henry
;                  Lal, Preeti
;                  Shah, Purvi
;                  Corley, Neil C.
;       TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                           PROLIFERATION
;       NUMBER OF SEQUENCES: 9
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Incyte Pharmaceuticals, Inc.
;            STREET: 3174 Porter Dr.
;            CITY: Palo Alto
;            STATE: CA
;            COUNTRY: USA
;            ZIP: 94304
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Diskette
;            COMPUTER: IBM Compatible
;            OPERATING SYSTEM: DOS
;            SOFTWARE: FastSEQ for Windows Version 2.0
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/894,657
;            FILING DATE: 28-Jun-2001
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: 09/410,372
;            FILING DATE: <Unknown>
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Billings, Lucy J.
;            REGISTRATION NUMBER: 36,749
;            REFERENCE/DOCKET NUMBER: PF-0421 US
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: 650-855-0555
;            TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 168 amino acids
;            TYPE: amino acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       IMMEDIATE SOURCE:
;            LIBRARY: GenBank
;            CLONE: 1683637
;       SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match             86.9%;  Score 73;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        108 QRYGRELRRMSDEF 121


RESULT 7
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match             86.9%;  Score 73;  DB 14;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00023;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        108 QRYGRELRRMSDEF 121


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECIF
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHIN
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match             79.8%;  Score 67;  DB 15;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 0.00017;
  Matches   13;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              |||||||||||||
Db          1 GRELRRMSDEFEG 13


RESULT 9
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
```

```
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

  Query Match            59.5%;  Score 50;  DB 15;  Length 215;
  Best Local Similarity  56.2%;  Pred. No. 1.6;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              :|:| :|||| || :|
Db        108 ERWGGDLRRMRDEADG 123


RESULT 10
US-09-843-846-18
; Sequence 18, Application US/09843846
; Patent No. US20020042362A1
;   GENERAL INFORMATION:
;        APPLICANT: KUNSCH, CHARLES A
;                   CHOPRA, ARVIND
;                   ROSEN, CRAIG A
;        TITLE OF INVENTION: HUMAN B-CELL TRANSLOCATION GENES-2 AND 3
;                            ANTIBODIES
;        NUMBER OF SEQUENCES: 18
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
;             STREET: 1100 NEW YORK AVENUE, NW, SUITE 600
;             CITY: WASHINGTON
;             STATE: DC
;             COUNTRY: USA
;             ZIP: 20005-3934
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/843,846
;             FILING DATE: 30-Apr-2001
;             CLASSIFICATION: <Unknown>
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: US 09/221,844
;             FILING DATE: 29-DEC-1998
;             APPLICATION NUMBER: US 08/718,738
;             FILING DATE: 18-SEP-1996
;             APPLICATION NUMBER: US 08/463,382
;             FILING DATE: 05-JUN-1995
;             APPLICATION NUMBER: US 08/460,104
;             FILING DATE: 02-JUN-1995
;             APPLICATION NUMBER: PCT/US95/03323
;             FILING DATE: 17-MAR-1995
;        ATTORNEY/AGENT INFORMATION:
;             NAME: STEFFE, ERIC K.
;             REGISTRATION NUMBER: 36,688
;             REFERENCE/DOCKET NUMBER: 1488.0730005/EKS/PSC
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: (202) 371-2600
;             TELEFAX: (202) 371-2540
;   INFORMATION FOR SEQ ID NO: 18:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 213 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: not relevant
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-843-846-18

  Query Match            51.2%;  Score 43;  DB 9;  Length 213;
  Best Local Similarity  43.8%;  Pred. No. 21;
  Matches    7;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              :|:| || |:  ::||
Db         22 RRFGEELERLLKKYEG 37


RESULT 11
US-09-331-631A-22
; Sequence 22, Application US/09331631A
; Patent No. US20020168392A1
; GENERAL INFORMATION:
;  APPLICANT:  Manners, John M.
;  APPLICANT:  Marcus, John Paul
;  APPLICANT:  Goulter, Kenneth C.
;  APPLICANT:  Green, Jodie L.
;  TITLE OF INVENTION:  ANTIMICROBIAL PROTEINS
;  FILE REFERENCE:  CULLN23.001APC
;  CURRENT APPLICATION NUMBER: US/09/331,631A
;  CURRENT FILING DATE:  1999-06-21
;  PRIOR APPLICATION NUMBER:  PCT/AU97/00874
;  PRIOR FILING DATE:  1997-12-22
;  PRIOR APPLICATION NUMBER:  AU PO 4275
;  PRIOR FILING DATE:  1996-12-20
;  NUMBER OF SEQ ID NOS:  40
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 22
;   LENGTH: 582
;   TYPE: PRT
;   ORGANISM: Maize
US-09-331-631A-22

  Query Match            48.8%;  Score 41;  DB 10;  Length 582;
  Best Local Similarity  50.0%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              :|:||| |   :| ||
Db        533 ERHGREEREKEEEREG 548


RESULT 12
US-10-156-761-13594
; Sequence 13594, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 13594
;   LENGTH: 380
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-13594

  Query Match            47.6%;  Score 40;  DB 15;  Length 380;
  Best Local Similarity  63.6%;  Pred. No. 1.2e+02;
  Matches    7;  Conservative    3;  Mismatches    1;  Indels    0;  Gaps    0;
```

```
Qy          3 YGRELRRMSDE 13
              | :||||::||
Db        241 YEKELRRLADE 251


RESULT 13
US-10-146-473-41
; Sequence 41, Application US/10146473
; Publication No. US20030108888A1
; GENERAL INFORMATION:
;  APPLICANT: Scanlan, Matthew
;  APPLICANT:  Gout, Ivan
;  APPLICANT:  Stockert, Elisabeth
;  APPLICANT:  Gure, Ali
;  APPLICANT:  Chen, Yao-Tseng
;  APPLICANT:  Old, Lloyd
;  TITLE OF INVENTION: Breast Cancer Antigens
;  FILE REFERENCE: L00461/70130(JRV)
;  CURRENT APPLICATION NUMBER: US/10/146,473
;  CURRENT FILING DATE:  2002-05-15
;  PRIOR APPLICATION NUMBER: US 60/291,150
;  PRIOR FILING DATE: 2001-05-15
;  NUMBER OF SEQ ID NOS: 82
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41
;   LENGTH: 2871
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-146-473-41

  Query Match             47.6%;  Score 40;  DB 15;  Length 2871;
  Best Local Similarity   60.0%;  Pred. No. 9.7e+02;
  Matches    9;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFE 15
              ||   ||||:| | |
Db       1648 QRTQEELRRLSSEVE 1662


RESULT 14
US-10-166-087-16
; Sequence 16, Application US/10166087
; Publication No. US20030077767A1
; GENERAL INFORMATION:
;  APPLICANT: Ecopia Biosciences Inc.
;  APPLICANT:  Farnet, Chris
;  APPLICANT:  Staffa, Alfredo
;  APPLICANT:  Zazopoulos, Emmanuel
;  TITLE OF INVENTION: Genes and proteins for the biosynthesis of anthramycin
;  FILE REFERENCE: 3014-2US
;  CURRENT APPLICATION NUMBER: US/10/166,087
;  CURRENT FILING DATE:  2002-06-11
;  NUMBER OF SEQ ID NOS: 51
;  SOFTWARE: PatentIn version 3.0
; SEQ ID NO 16
;   LENGTH: 764
;   TYPE: PRT
;   ORGANISM: Streptomyces refuineus subspecies thermotolerans
US-10-166-087-16

  Query Match             45.8%;  Score 38.5;  DB 15;  Length 764;
  Best Local Similarity   62.5%;  Pred. No. 4.3e+02;
  Matches   10;  Conservative    2;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 QRYGRELRRMSDEFEG 16
              || ||||   |:|:||
Db        220 QRPGREL-TYSNEYEG 234


RESULT 15
US-10-092-750-1
; Sequence 1, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
;  APPLICANT: Hammond, Philip W.
;  APPLICANT:  Alpin, Julia
;  APPLICANT:  Wright, Martin C.
;  TITLE OF INVENTION: Polypeptides Interactive with BCL-Xl
;  FILE REFERENCE: 50036/050002
;  CURRENT APPLICATION NUMBER: US/10/092,750
;  CURRENT FILING DATE:  2002-03-07
;  PRIOR APPLICATION NUMBER: US 60/274,526
;  PRIOR FILING DATE: 2001-03-08
;  NUMBER OF SEQ ID NOS: 253
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 1
;   LENGTH: 35
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-092-750-1

  Query Match             45.2%;  Score 38;  DB 15;  Length 35;
  Best Local Similarity   70.0%;  Pred. No. 21;
  Matches    7;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 RELRRMSDEF 14
              :||||: |||
Db         19 QELRRIGDEF 28


Search completed: September 15, 2003, 17:47:52
Job time : 12.6857 secs
```

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 7.2 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	100.0	204	2	A55671	bad protein - mous
2	44	52.4	564	2	H75403	glycosyl hydrolase
3	43	51.2	463	2	H83036	probable two-compo
4	42.5	50.6	1014	2	T36031	excinuclease ABC c
5	42	50.0	370	2	S38185	2-dehydro-3-deoxy-
6	42	50.0	445	2	G97123	probable Fe-S oxid
7	41	48.8	84	2	F84388	hypothetical prote
8	41	48.8	113	2	JQ1128	gas-vesicle operon
9	41	48.8	113	2	T08234	gas-vesicle operon
10	41	48.8	220	2	F72289	oxidoreductase, so
11	41	48.8	271	2	AB1762	phosphate ABC tran
12	41	48.8	271	2	AH1386	phosphate ABC tran
13	41	48.8	322	1	RGBYC3	regulatory protein
14	41	48.8	398	2	G69496	ATP-dependent 26S
15	41	48.8	582	2	B53234	vicilin-like stora
16	40.5	48.2	334	2	A39172	Antho-RFamide neur
17	40.5	48.2	435	2	A44308	Antho-RFamide prec
18	40	47.6	206	2	C36365	transforming prote
19	40	47.6	219	2	T15466	hypothetical prote
20	40	47.6	374	2	C84338	spermidine/putresc
21	40	47.6	411	2	F87644	transcription regu
22	40	47.6	887	2	T38885	probable ATP-depen
23	40	47.6	967	2	F82668	oxoglutarate dehyd
24	40	47.6	2677	2	A38194	desmoplakin I - hu
25	40	47.6	5138	2	B96695	hypothetical prote
26	39.5	47.0	310	2	D86675	mevalonate kinase
27	39.5	47.0	910	2	G91024	NADH dehydrogenase
28	39.5	47.0	910	2	H85868	NADH dehydrogenase
29	39.5	47.0	910	2	A65000	NADH2 dehydrogenas
30	39.5	47.0	910	2	AI0796	NADH2 dehydrogenas
31	39.5	47.0	914	2	AD0311	NADH2 dehydrogenas
32	39	46.4	73	2	AC3365	hypothetical prote
33	39	46.4	87	2	A82928	hypothetical prote
34	39	46.4	91	2	T26996	ribosomal protein
35	39	46.4	105	2	B84184	hypothetical prote
36	39	46.4	232	2	A42095	floral homeotic pr
37	39	46.4	275	2	C69808	transporter homolo
38	39	46.4	315	2	T39444	hypothetical prote
39	39	46.4	331	2	E90121	DNA repair protein
40	39	46.4	365	2	S42107	RAD51 protein homo
41	39	46.4	380	2	T32163	hypothetical prote
42	39	46.4	501	2	H84727	probable inositol
43	39	46.4	574	2	T48113	inositol-1,4,5-tri
44	39	46.4	1016	2	H75356	excinuclease ABC c
45	39	46.4	1248	2	G83278	cobalamin biosynth

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 100.0%; Score 84; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFEG 16
||||||||||||||||
Db 145 QRYGRELRRMSDEFEG 160

RESULT 2
H75403
glycosyl hydrolase, family 13 - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: H75403
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <WHI>
A;Cross-references: GB:AE001983; GB:AE000513; NID:g6459123; PIDN:AAF10944.1; PID:g645
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1375
A;Map position: 1
C;Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 52.4%; Score 44; DB 2; Length 564;
Best Local Similarity 61.5%; Pred. No. 23;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```
Qy          3 YGRELRRMSDEFE 15
              | ||:||: |||:
Db        283 YVREMRRVIDEFD 295


RESULT 3
H83036
probable two-component sensor PA4886 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 14-Sep-2001
C;Accession: H83036
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83036
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-463 <STO>
A;Cross-references: GB:AE004901; GB:AE004091; NID:g9951147; PIDN:AAG08271.1; GSPDB:GN001
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4886
C;Superfamily: hypothetical protein HI1707; sensor histidine kinase homology

  Query Match             51.2%;  Score 43;  DB 2;  Length 463;
  Best Local Similarity   61.5%;  Pred. No. 28;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              | |||| :: |||
Db        321 GEELRRQAEYFEG 333


RESULT 4
T36031
excinuclease ABC chain A SCC54.18c [similarity] - Streptomyces coelicolor
N;Contains: excision endonuclease ABC (EC 3.1.-.-) chain A
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Feb-2001
C;Accession: T36031
R;Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, March 1999
A;Reference number: Z21581
A;Accession: T36031
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-1014 <SEE>
A;Cross-references: EMBL:AL035591; PIDN:CAB38148.1; GSPDB:GN00070; SCOEDB:SCC54.18c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: uvrA; SCOEDB:SCC54.18c
C;Superfamily: excinuclease ABC chain A; ATP-binding cassette homology
C;Keywords: ATP; DNA binding; DNA repair; hydrolase; nucleotide binding; P-loop
F;32-39/Region: nucleotide-binding motif A (P-loop)
F;645-652/Region: nucleotide-binding motif A (P-loop)

  Query Match             50.6%;  Score 42.5;  DB 2;  Length 1014;
  Best Local Similarity   66.7%;  Pred. No. 75;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 RYGRELRRMSDEFEG 16
              ||||| || :  |||
Db        368 RYGRE-RRYTTAFEG 381


RESULT 5
S38185
2-dehydro-3-deoxy-phosphoheptonate aldolase (EC 4.1.2.15) ARO4 - yeast (Saccharomyces ce
N;Alternate names: 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase; DAHP synthase;
C;Species: Saccharomyces cerevisiae
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 03-Jun-2002
C;Accession: S38185; S46126; S46130; JN0322; B48651
R;Doignon, F.; Biteau, N.; Aigle, M.; Crouzet, M.
Yeast 9, 1131-1137, 1993
A;Title: The complete sequence of a 6794 bp segment located on the right arm of chrom
A;Reference number: S38185; MUID:94078675; PMID:8256522
A;Accession: S38185
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-370 <DOI>
A;Cross-references: GB:L20296; NID:g311101; PIDN:AAA65607.1; PID:g311102
R;Aljinovic, G.; Pohl, F.M.; Pohl, T.M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45906
A;Accession: S46126
A;Molecule type: DNA
A;Residues: 1-370 <ALJ>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Aigle, M.; Baclet, M.C.; Barthe, C.; Biteau, N.; Crouzet, M.; Doignon, F.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45940
A;Accession: S46130
A;Molecule type: DNA
A;Residues: 1-370 <AIG>
A;Cross-references: EMBL:Z36118; NID:g536664; PIDN:CAA85212.1; PID:g536665; MIPS:YBR2
R;Kuenzler, M.; Paravicini, G.; Egli, C.M.; Irniger, S.; Braus, G.H.
Gene 113, 67-74, 1992
A;Title: Cloning, primary structure and regulation of the ARO4 gene, encoding the tyr
A;Reference number: JN0322; MUID:92225349; PMID:1348717
A;Accession: JN0322
A;Molecule type: DNA
A;Residues: 1-204,208-370 <KUE>
A;Cross-references: EMBL:X61107
R;Kuenzler, M.; Balmelli, T.; Egli, C.M.; Paravicini, G.; Braus, G.H.
J. Bacteriol. 175, 5548-5558, 1993
A;Title: Cloning, primary structure, and regulation of the HIS7 gene encoding a bifun
A;Reference number: A48651; MUID:93374850; PMID:8366040
A;Accession: B48651
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 352-370 <KU2>
A;Cross-references: GB:X61107
C;Comment: This enzyme catalyzes the condensation of phosphoenolpyruvate and D-erythr
C;Genetics:
A;Gene: SGD:ARO4
A;Cross-references: SGD:S0000453; MIPS:YBR249c
A;Map position: 2R
C;Function:
A;Description: aldehyde-lyase; carbon-carbon lyase
A;Pathway: aromatic amino acid biosynthesis; shikimate pathway
A;Note: first step in shikimate pathway
C;Superfamily: phospho-2-dehydro-3-deoxyheptonate aldolase
C;Keywords: aldehyde-lyase; aromatic amino acid biosynthesis; carbon-carbon lyase; cy

  Query Match             50.0%;  Score 42;  DB 2;  Length 370;
  Best Local Similarity   43.8%;  Pred. No. 32;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              | |   |:::||| :|
Db         85 QEYALRLKKLSDELKG 100


RESULT 6
G97123
probable Fe-S oxidoreductase CAC1813 [imported] - Clostridium acetobutylicum
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 30-Sep-2001
C;Accession: G97123
R;Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
```

A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: G97123
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-445 <KUR>
A;Cross-references: GB:AE001437; PIDN:AAK79778.1; PID:g15024787; GSPDB:GN00168
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC1813
C;Superfamily: conserved hypothetical protein b0835

Query Match 50.0%; Score 42; DB 2; Length 445;
Best Local Similarity 46.7%; Pred. No. 39;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFE 15
||||:||: : | :
Db 85 QRYGKELKELVPEID 99

RESULT 7
F84388
hypothetical protein Vng2379h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84388
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84 <STO>
A;Cross-references: GB:AE004437; NID:g10581786; PIDN:AAG20474.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG2379H

Query Match 48.8%; Score 41; DB 2; Length 84;
Best Local Similarity 57.1%; Pred. No. 10;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 RYGRELRRMSDEFE 15
||| | || || :
Db 66 RYGTEASRMRDEHD 79

RESULT 8
JQ1128
gas-vesicle operon protein gvpK - Halobacterium salinarum plasmids pHH1 and pNRC100
C;Species: Halobacterium salinarum
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 31-Jan-2000
C;Accession: JQ1128; S15190
R;Jones, J.G.; Young, D.C.; DasSarma, S.
Gene 102, 117-122, 1991
A;Title: Structure and organization of the gas vesicle gene cluster on the Halobacterium
A;Reference number: JQ1122; MUID:91323716; PMID:1864501
A;Accession: JQ1128
A;Molecule type: DNA
A;Residues: 1-113 <JON>
A;Cross-references: GB:M58557; NID:g150406; PID:g455299
A;Experimental source: strain NRC-1, plasmid pNRC100
A;Genetics: NRC
A;Note: the source is designated as Halobacterium halobium
R;Horne, M.; Englert, C.; Wimmer, C.; Pfeifer, F.
Mol. Microbiol. 5, 1159-1174, 1991
A;Title: A DNA region of 9 kbp contains all genes necessary for gas vesicle synthesis in
A;Reference number: S15183; MUID:92065812; PMID:1956294
A;Accession: S15190
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-113 <HOR>
A;Cross-references: EMBL:X55648; NID:g43516; PID:g43524
A;Experimental source: plasmid pHH1
A;Genetics: PHH
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, September 19
A;Note: the source is designated as Halobacterium halobium
C;Genetics: <NRC>
A;Gene: gvpK
A;Genome: plasmid pNRC100
C;Genetics: <PHH>
A;Gene: gvpK
A;Genome: plasmid pHH1

Query Match 48.8%; Score 41; DB 2; Length 113;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFE 15
:| ||:|: : || |
Db 50 ERLGRQLQALEDELE 64

RESULT 9
T08234
gas-vesicle operon protein gvpK - Halobacterium sp. (strain NRC-1) plasmid *pNRC100*
N;Alternate names: hypothetical protein H0228
C;Species: Halobacterium sp.
A;Variety: strain NRC-1
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 31-Jan-2000
C;Accession: T08234
R;Ng, W.V.; Ciufo, S.A.; Smith, T.M.; Bumgarner, R.E.; Baskin, D.; Faust, J.; Hall,
Genome Res. 8, 1131-1141, 1998
A;Title: Snapshot of a large dynamic replicon in a halophilic Archaeon: megaplasmid
A;Reference number: Z16408; MUID:99063795; PMID:9847077
A;Accession: T08234
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-113 <NGW>
A;Cross-references: EMBL:AF016485; NID:g2822278; PID:g2822295; HALOSP:H0228
A;Experimental source: strain NRC-1
C;Genetics:
A;Gene: gvpK; HALOSP:H0228
A;Genome: plasmid pNRC100

Query Match 48.8%; Score 41; DB 2; Length 113;
Best Local Similarity 46.7%; Pred. No. 14;
Matches 7; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFE 15
:| ||:|: : || |
Db 50 ERLGRQLQALEDELE 64

RESULT 10
F72289
oxidoreductase, sol/devB family - Thermotoga maritima (strain MSB8)
C;Species: Thermotoga maritima
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C;Accession: F72289
R;Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hi
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
C.M.
Nature 399, 323-329, 1999
A;Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
A;Reference number: A72200; MUID:99287316; PMID:10360571
A;Accession: F72289
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-220 <ARN>
A;Cross-references: GB:AE001772; GB:AE000512; NID:g4981693; PIDN:AAD36230.1; PID:g49
A;Experimental source: strain MSB8

```
C;Genetics:
A;Gene: TM1154
C;Superfamily: yeast SOL3 protein

  Query Match             48.8%;  Score 41;  DB 2;  Length 220;
  Best Local Similarity   40.0%;  Pred. No. 27;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFE 15
              ::| ||:|  :|:|:
Db        113 EKYEREIRSATDQFD 127


RESULT 11
AB1762
phosphate ABC transporter (ATP-binding protein) homolog lin2639 [imported] - Listeria in
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C;Accession: AB1762
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1762
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-271 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97866.1; PID:g16415176; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2639
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology

  Query Match             48.8%;  Score 41;  DB 2;  Length 271;
  Best Local Similarity   57.1%;  Pred. No. 34;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFEG 16
              | | | ||:|| :|
Db         66 YLRALNRMNDEIDG 79


RESULT 12
AH1386
phosphate ABC transporter (ATP-binding protein) homolog lmo2496 [imported] - Listeria mo
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C;Accession: AH1386
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker
.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
D.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A;Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AH1386
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-271 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00574.1; PID:g16411984; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2496
C;Superfamily: inner membrane protein malK; ATP-binding cassette homology

  Query Match             48.8%;  Score 41;  DB 2;  Length 271;
  Best Local Similarity   57.1%;  Pred. No. 34;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFEG 16
              | | | ||:|| :|
Db         66 YLRALNRMNDEIDG 79


RESULT 13
RGBYC3
regulatory protein SNF4 - yeast (Saccharomyces cerevisiae)
N;Alternate names: CAT3 protein; protein G2945; protein YGL115w
C;Species: Saccharomyces cerevisiae
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jul-2000
C;Accession: A38906; JT0316; S48508; S64125
R;Celenza, J.L.; Eng, F.J.; Carlson, M.
Mol. Cell. Biol. 9, 5045-5054, 1989
A;Title: Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae: evidence f
A;Reference number: A33480; MUID:90097921; PMID:2481228
A;Accession: A38906
A;Molecule type: DNA
A;Residues: 1-322 <CEL>
A;Cross-references: GB:M30470; NID:g172635; PIDN:AAA35061.1; PID:g172636
R;Schueller, H.J.; Entian, K.D.
Gene 67, 247-257, 1988
A;Title: Molecular characterization of yeast regulatory gene CAT3 necessary for gluc
A;Reference number: JT0316; MUID:89006284; PMID:3049255
A;Accession: JT0316
A;Molecule type: DNA
A;Residues: 1-322 <SCH>
A;Cross-references: GB:M21760; NID:g171164; PIDN:AAA34472.1; PID:g171165
R;Doi, A.; Doi, K.
submitted to the EMBL Data Library, June 1993
A;Description: Correct end of the ORF for the CDC20 gene of Saccharomyces cerevisiae
A;Reference number: S48507
A;Accession: S48508
A;Molecule type: DNA
A;Residues: 1-21 <DOI>
A;Cross-references: EMBL:D16506; NID:g391938; PIDN:BAA03958.1; PID:g2160324
R;Lauquin, G.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64122
A;Accession: S64125
A;Molecule type: DNA
A;Residues: 1-322 <LAU>
A;Cross-references: EMBL:Z72637; NID:g1322666; PIDN:CAA96823.1; PID:g1322667; GSPDB:(
A;Experimental source: strain S288C
C;Genetics:
A;Gene: SGD:SNF4; CAT34; MIPS:YGL115w
A;Cross-references: SGD:S0003083; MIPS:YGL115w
A;Map position: 7L
C;Function:
A;Description: involved in derepression of glucose-repressed genes
C;Superfamily: CAT3 protein
C;Keywords: nucleus; transcription regulation

  Query Match             48.8%;  Score 41;  DB 1;  Length 322;
  Best Local Similarity   61.5%;  Pred. No. 41;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              |  | | ||:|||
Db        258 GEALMRRSDDFEG 270


RESULT 14
G69496
ATP-dependent 26S proteinase regulatory subunit 4 homolog - Archaeoglobus fulgidus
C;Species: Archaeoglobus fulgidus
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 19-Jan-2001
C;Accession: G69496
R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Doc
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
```

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaec
A;Reference number: A69250; MUID:98049343; PMID:9389475
A;Accession: G69496
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-398 <KLE>
A;Cross-references: GB:AE000967; GB:AE000782; NID:g2689290; PIDN:AAB89280.1; PID:g264856
C;Superfamily: ATP-dependent 26S proteinase; FtsH/SEC18/CDC48-type ATP-binding domain hc
C;Keywords: ATP; nucleotide binding; P-loop
F;155-365/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VATP>
F;182-189/Region: nucleotide-binding motif A (P-loop)

  Query Match 48.8%; Score 41; DB 2; Length 398;
  Best Local Similarity 57.1%; Pred. No. 51;
  Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

```
Qy          2 RYGRELRRMSDEFE 15
              || ||:||:  | |
Db         43 RYEREVRRLRSEVE 56
```

RESULT 15
B53234
vicilin-like storage protein Glb1-L, embryo - maize
N;Alternate names: globulin-1L
C;Species: Zea mays (maize)
C;Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 24-Nov-1999
C;Accession: B53234; S21824
R;Belanger, F.C.; Kriz, A.L.
Genetics 129, 863-872, 1991
A;Title: Molecular basis for allelic polymorphism of the maize Globulin-1 gene.
A;Reference number: A53234; MUID:92090707; PMID:1752424
A;Accession: B53234
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-582 <BEL>
A;Cross-references: EMBL:X59083; NID:g22283; PIDN:CAA41809.1; PID:g22284
A;Experimental source: inbred line W64A6
A;Note: sequence extracted from NCBI backbone (NCBIP:71285)
C;Genetics:
A;Gene: Glb1-L
A;Introns: 167/1; 225/3; 252/3; 349/3
C;Superfamily: glycinin

  Query Match 48.8%; Score 41; DB 2; Length 582;
  Best Local Similarity 50.0%; Pred. No. 75;
  Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEFEG 16
              :|:||| |   :| ||
Db        533 ERHGREEREKEEEREG 548
```

Search completed: September 15, 2003, 17:27:01
Job time : 7.2 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 3.77143 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	84	100.0	204	1	BAD_MOUSE	Q61337 mus musculu
2	84	100.0	205	1	BAD_RAT	O35147 rattus norv
3	73	86.9	168	1	BAD_HUMAN	Q92934 homo sapien
4	42.5	50.6	1014	1	UVRA_STRCO	Q9z507 streptomyce
5	42	50.0	196	1	BIM_MOUSE	O54918 mus musculu
6	42	50.0	196	1	BIM_RAT	O88498 rattus norv
7	42	50.0	370	1	AROG_YEAST	P32449 saccharomyc
8	42	50.0	653	1	HT2A_HUMAN	Q13049 homo sapien
9	41	48.8	113	1	GVK1_HALN1	P24375 halobacteri
10	41	48.8	220	1	6PGL_THEMA	Q9x0n8 thermotoga
11	41	48.8	322	1	SNF4_YEAST	P12904 saccharomyc
12	41	48.8	328	1	SNF4_KLULA	Q9p869 kluyveromyc
13	41	48.8	398	1	PSMR_ARCFU	O28303 archaeoglob
14	40.5	48.2	334	1	FMRA_CALPA	Q01133 calliactis
15	40.5	48.2	429	1	FMR2_ANTEL	Q16994 anthopleura
16	40.5	48.2	435	1	FMR1_ANTEL	P10419 anthopleura
17	40	47.6	205	1	RAS3_RHIRA	P22280 rhizomucor
18	40	47.6	887	1	YAY5_SCHPO	Q10213 schizosacch
19	40	47.6	2871	1	DESP_HUMAN	P15924 homo sapien
20	39.5	47.0	907	1	NUOG_ECOLI	P33602 escherichia
21	39.5	47.0	907	1	NUOG_SALTY	P33900 salmonella
22	39	46.4	87	1	Y152_UREPA	Q9pqz2 ureaplasma
23	39	46.4	95	1	R37A_HALN1	Q9hsg8 halobacteri
24	39	46.4	232	1	AP3_ARATH	P35632 arabidopsis
25	39	46.4	365	1	RA51_SCHPO	P36601 schizosacch
26	39	46.4	889	1	SEC3_DROME	Q9vvg4 drosophila
27	39	46.4	1016	1	UVRA_DEIRA	Q46577 deinococcus
28	39	46.4	1535	1	LML1_CAEEL	Q18823 caenorhabdi
29	39	46.4	8797	1	SNE1_HUMAN	Q8nf91 homo sapien
30	38.5	45.8	323	1	FRG1_CAEEL	O18282 caenorhabdi
31	38.5	45.8	468	1	SELA_PSEAE	Q9hv01 pseudomonas
32	38	45.2	177	1	IF3_CLOPE	Q8xj67 clostridium
33	38	45.2	198	1	BIM_HUMAN	O43521 homo sapien
34	38	45.2	207	1	THIE_PYRAB	Q9uzq5 pyrococcus
35	38	45.2	370	1	AROG_CANAL	P79023 candida alb
36	38	45.2	570	1	STIM_DROME	P83094 drosophila
37	38	45.2	712	1	PERL_HUMAN	P22079 homo sapien
38	38	45.2	861	1	GCR3_YEAST	P34160 saccharomyc
39	38	45.2	1047	1	RIR1_CHLMU	Q9pl93 chlamydia m
40	38	45.2	1047	1	RIR1_CHLTR	O84834 chlamydia t
41	38	45.2	1268	1	VGLN_HUMAN	Q00341 homo sapien
42	38	45.2	1270	1	VGLN_CHICK	P81021 gallus gall
43	38	45.2	1958	1	UBR1_SCHPO	O60152 schizosacch
44	37.5	44.6	992	1	UVRA_MICLU	P13567 micrococcus
45	37	44.0	126	1	LEU3_BUCUL	Q9aqc8 buchnera ap

ALIGNMENTS

RESULT 1
BAD_MOUSE
ID BAD_MOUSE STANDARD; PRT; 204 AA.
AC Q61337;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE 6) (Bcl-xL/Bcl-2 associated death promoter).
GN BAD OR BBC6.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Thymus;
RX MEDLINE=95136361; PubMed=7834748;
RA Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT promotes cell death.";
RL Cell 80:285-291(1995).
RN [2]
RP PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX MEDLINE=98022383; PubMed=9381178;
RA Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT "Interleukin-3-induced phosphorylation of BAD through the protein
RT kinase Akt.";
RL Science 278:687-689(1997).
RN [3]
RP MUTAGENESIS OF SERINE RESIDUES.
RX MEDLINE=20403302; PubMed=10949026;
RA Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA Greenberg M.E.;
RT "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT BH3 domain phosphorylation.";
RL Mol. Cell 6:41-51(2000).
CC -!- FUNCTION: Promotes cell death. Successfully competes for the
CC binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC of heterodimerization of these proteins with BAX. Can reverse the
CC death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC Appears to act as a link between growth factor receptor signaling
CC and the apoptotic pathways.
CC -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC phosphorylation, locates to the cytoplasm.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC BAX for their pro-apoptotic activity and for their interaction
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC with 14-3-3 proteins. This interaction then facilitates the
CC phosphorylation at Ser-155, a site within the BH3 domain, leading
CC to the release of Bcl-x(L) and the promotion of cell survival.

```
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             100.0%;  Score 84;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        145 QRYGRELRRMSDEFEG 160


RESULT 2
BAD_RAT
ID   BAD_RAT        STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;
```

```
  Query Match             100.0%;  Score 84;  DB 1;  Length 205;
  Best Local Similarity   100.0%;  Pred. No. 1.3e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              ||||||||||||||||
Db        146 QRYGRELRRMSDEFEG 161


RESULT 3
BAD_HUMAN
ID   BAD_HUMAN      STANDARD;      PRT;   168 AA.
AC   Q92934; O14803;
DT   01-NOV-1997 (Rel. 35, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE   XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN   BAD OR BBC6 OR BCL2L8.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT   "A human protein that interacts with Bcl-2 and have homology to mouse
RT   BAD.";
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX   MEDLINE=97083574; PubMed=8929532;
RA   Wang H.-G., Rapp U.R., Reed J.C.;
RT   "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL   Cell 87:629-638(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Takayama S., Reed J.C.;
RL   Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A., AND DIMERIZATION.
RC   TISSUE=Bone marrow;
RX   MEDLINE=98049554; PubMed=9388232;
RA   Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA   Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT   "Dimerization properties of human BAD.";
RL   J. Biol. Chem. 272:30866-30872(1997).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Lung;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match             86.9%;  Score 73;  DB 1;  Length 168;
```

```
  Best Local Similarity   100.0%;  Pred. No. 6.7e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        108 QRYGRELRRMSDEF 121


RESULT 4
UVRA_STRCO
ID   UVRA_STRCO     STANDARD;      PRT;  1014 AA.
AC   Q9Z507;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   UvrABC system protein A (UvrA protein) (Excinuclease ABC subunit A).
GN   UVRA OR SCO1958 OR SCC54.18C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
CC   -!- FUNCTION: The UvrABC repair system catalyzes the recognition and
CC       processing of DNA lesions. UvrA is an ATPase and a DNA-binding
CC       protein. A damage recognition complex composed of 2 uvrA and 2
CC       uvrB subunits scans DNA for abnormalities. When the presence of a
CC       lesion has been verified by uvrB, the uvrA molecules dissociate
CC       (By similarity).
CC   -!- SUBUNIT: Forms a heterotetramer with uvrB during the search for
CC       lesions (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AL035591; CAB38148.1; -.
DR   PIR; T36031; T36031.
DR   HAMAP; MF_00205; -; 1.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR004602; UvrA.
DR   Pfam; PF00005; ABC_tran; 2.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   TIGRFAMs; TIGR00630; uvra; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER_1; 2.
DR   PROSITE; PS50893; ABC_TRANSPORTER_2; 1.
KW   SOS response; Excision nuclease; DNA repair; DNA recombination;
KW   DNA excision; ATP-binding; DNA-binding; Repeat; Zinc; Metal-binding;
KW   Zinc-finger; Complete proteome.
FT   NP_BIND      32     39       ATP (POTENTIAL).
FT   NP_BIND     645    652       ATP (POTENTIAL).
FT   ZN_FING     744    770       C4-TYPE.
SQ   SEQUENCE   1014 AA;  110997 MW;  084D6B18692A792D CRC64;

  Query Match           50.6%;  Score 42.5;  DB 1;  Length 1014;
  Best Local Similarity   66.7%;  Pred. No. 39;
  Matches   10;  Conservative    1;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 RYGRELRRMSDEFEG 16
              ||||| || :  |||
Db        368 RYGRE-RRYTTAFEG 381


RESULT 5
BIM_MOUSE
ID   BIM_MOUSE      STANDARD;      PRT;   196 AA.
AC   O54918; O54919; O54920;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BCL2-like protein 11 (BCL2 interacting mediator of cell death).
GN   BCL2L11 OR BIM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE
RP   SPECIFICITY, AND ALTERNATIVE SPLICING.
RX   MEDLINE=98094360; PubMed=9430630;
RA   O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,
RA   Cory S., Huang D.C.S.;
RT   "Bim: a novel member of the Bcl-2 family that promotes apoptosis.";
RL   EMBO J. 17:384-395(1998).
CC   -!- FUNCTION: INDUCES APOPTOSIS. THE ISOFORMS VARY IN CYTOTOXICITY
CC       WITH ISOFORM BIMS BEING THE MOST POTENT AND ISOFORM BIMEL BEING
CC       THE LEAST POTENT.
CC   -!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
CC       PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
CC       NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
CC       BAX OR BAK (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOPLASMIC MEMBRANES.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=BimEL;
CC         IsoId=O54918-1; Sequence=Displayed;
CC       Name=BimL;
CC         IsoId=O54918-2; Sequence=VSP_000536;
CC       Name=BimS;
CC         IsoId=O54918-3; Sequence=VSP_000537;
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF B-AND T-LYMPHOID CELL
CC       LINES.
CC   -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
CC       CYTOTOXICITY.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF032459; AAC40029.1; -.
DR   EMBL; AF032460; AAC40030.1; -.
DR   EMBL; AF032461; AAC40031.1; -.
DR   MGD; MGI:1197519; Bcl2l11.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Alternative splicing; Membrane.
FT   DOMAIN      146    160       BH3.
FT   VARSPLIC     42     97       Missing (in isoform BimL).
FT                                /FTId=VSP_000536.
FT   VARSPLIC     42    127       Missing (in isoform BimS).
FT                                /FTId=VSP_000537.
```

```
SQ   SEQUENCE   196 AA;  22066 MW;  531C176E5F1AC9AA CRC64;

  Query Match             50.0%;  Score 42;  DB 1;  Length 196;
  Best Local Similarity   61.5%;  Pred. No. 8.6;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEF 14
              |  :||||: |||
Db        145 RIAQELRRIGDEF 157


RESULT 6
BIM_RAT
ID   BIM_RAT        STANDARD;      PRT;   196 AA.
AC   O88498; O88497; Q9WUI8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   BCL2-like protein 11 (BCL2 interacting mediator of cell death)
DE   (Bcl-2 related ovarian death protein).
GN   BCL2L11 OR BIM OR BOD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND TISSUE SPECIFICITY
RP   (ISOFORMS BOD-L; BOD-M AND BOD-S).
RC   TISSUE=Ovary;
RX   MEDLINE=98400436; PubMed=9731710;
RA   Hsu S.Y., Lin P., Hsueh A.J.W.;
RT   "BOD (Bcl-2-related ovarian death gene) is an ovarian BH3 domain-
RT   containing proapoptotic Bcl-2 protein capable of dimerization with
RT   diverse antiapoptotic Bcl-2 members.";
RL   Mol. Endocrinol. 12:1432-1440(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM BIML).
RA   Chen D., Simon R.P., Chen J.;
RT   "Cloning of rat bimEL and bimL, and their differential expression in
RT   ischemia and normal rat brain.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INDUCES APOPTOSIS.
CC   -!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
CC       PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
CC       NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
CC       BAX OR BAK.
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOPLASMIC MEMBRANES
CC       (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=BOD-L;
CC         IsoId=O88498-1; Sequence=Displayed;
CC         Note=Isoform BOD-S is produced by alternative initation at
CC         Met-104 of isoform BOD-L;
CC       Name=BimL;
CC         IsoId=O88498-2; Sequence=VSP_000538;
CC       Name=BOD-M;
CC         IsoId=O88498-3; Sequence=VSP_000539;
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, BOD-L (shown here) and BOD-S, are produced
CC         by alternative initiation at Met-1 and Met-104;
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
CC       CYTOTOXICITY.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF065433; AAC23595.1; -.
DR   EMBL; AF065431; AAC23593.1; -.
DR   EMBL; AF065432; AAC23594.1; -.
DR   EMBL; AF136927; AAD26594.1; -.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Alternative splicing; Membrane; Alternative initiation.
FT   CHAIN         1    196       BCL2-LIKE PROTEIN 11, ISOFORM BOD-L.
FT   CHAIN       104    196       BCL2-LIKE PROTEIN 11, ISOFORM BOD-S.
FT   INIT_MET    104    104       FOR ISOFORM BOD-S.
FT   DOMAIN      146    160       BH3.
FT   VARSPLIC     42     97       Missing (in isoform BimL).
FT                                /FTId=VSP_000538.
FT   VARSPLIC     42    127       Missing (in isoform BOD-M).
FT                                /FTId=VSP_000539.
FT   CONFLICT    136    136       E -> D (IN REF. 1; AAC23594).
SQ   SEQUENCE   196 AA;  22055 MW;  B4D2146F9C0B37A0 CRC64;

  Query Match             50.0%;  Score 42;  DB 1;  Length 196;
  Best Local Similarity   61.5%;  Pred. No. 8.6;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEF 14
              |  :||||: |||
Db        145 RIAQELRRIGDEF 157


RESULT 7
AROG_YEAST
ID   AROG_YEAST     STANDARD;      PRT;   370 AA.
AC   P32449;
DT   01-OCT-1993 (Rel. 27, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4 OR YBR249C OR YBR1701.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92225349; PubMed=1348717;
RA   Kuenzler M., Paravicini G., Egli C., Irniger S., Braus G.H.;
RT   "Cloning, primary structure and regulation of the ARO4 gene, encoding
RT   the tyrosine-inhibited 3-deoxy-D-arabino-heptulosonate-7-phosphate
RT   synthase from Saccharomyces cerevisiae.";
RL   Gene 113:67-74(1992).
RN   [2]
RP   REVISIONS TO 205-207.
RA   Kuenzler M.;
RL   Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RX   MEDLINE=94078675; PubMed=8256522;
RA   Doignon F., Biteau N., Aigle M., Crouzet M.;
RT   "The complete sequence of a 6794 bp segment located on the right arm
RT   of chromosome II of Saccharomyces cerevisiae. Finding of a putative
RT   dUTPase in a yeast.";
RL   Yeast 9:1131-1137(1993).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c;
RA   Aljinovic G., Pohl F.M., Pohl T.M.;
RL   Submitted (AUG-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
```

```
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE.
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- INDUCTION: By amino acid starvation.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X61107; CAA43419.1; -.
DR   EMBL; L20296; AAA65607.1; -.
DR   EMBL; Z36118; CAA85212.1; -.
DR   PIR; S38185; S38185.
DR   HSSP; P00886; 1QR7.
DR   SGD; S0000453; ARO4.
DR   GO; GO:0003849; F:2-dehydro-3-deoxyphosphoheptonate aldolase . . .; IDA.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  39749 MW;  594ED48F24175979 CRC64;

  Query Match            50.0%;  Score 42;  DB 1;  Length 370;
  Best Local Similarity  43.8%;  Pred. No. 17;
  Matches    7;  Conservative    4;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              | |   |:::||| :|
Db         85 QEYALRLKKLSDELKG 100


RESULT 8
HT2A_HUMAN
ID   HT2A_HUMAN     STANDARD;      PRT;   653 AA.
AC   Q13049; Q9NQP8;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc-finger protein HT2A (72 kDa Tat-interacting protein) (Tripartite
DE   motif-containing protein 32).
GN   TRIM32 OR HT2A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95297135; PubMed=7778269;
RA   Fridell R.A., Harding L.S., Bogerd H.P., Cullen B.R.;
RT   "Identification of a novel human zinc finger protein that
RT   specifically interacts with the activation domain of lentiviral Tat
RT   proteins.";
RL   Virology 209:347-357(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sehra H.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: MAY PLAY A SIGNIFICANT ROLE IN MEDIATING THE BIOLOGICAL
CC       ACTIVITY OF THE HIV-1 TAT PROTEIN IN VIVO. BINDS SPECIFICALLY TO
CC       THE ACTIVATION DOMAIN OF HIV-1 TAT AND CAN ALSO INTERACT WITH THE
CC       HIV-2 AND EIAV TAT PROTEINS IN VIVO.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: SPLEEN, THYMUS, PROSTRATE, TESTIS, OVARY,
CC       INTESTINE AND COLON.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC   -!- SIMILARITY: Contains 1 B box-type zinc finger.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U18543; AAA86474.1; -.
DR   EMBL; AL133284; CAB92723.1; -.
DR   EMBL; BC003154; AAH03154.1; -.
DR   HSSP; P29590; 1BOR.
DR   Genew; HGNC:16380; TRIM32.
DR   MIM; 602290; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   InterPro; IPR001258; NHL.
DR   InterPro; IPR000315; Znf_Bbox.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF01436; NHL; 5.
DR   Pfam; PF00643; zf-B_box; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00336; BBOX; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Zinc-finger; Nuclear protein.
FT   DOMAIN       2      6       POLY-ALA.
FT   ZN_FING     20     65       RING-TYPE.
FT   ZN_FING    103    133       B BOX-TYPE.
FT   CONFLICT    27     27       F -> I (IN REF. 1).
SQ   SEQUENCE   653 AA;  71988 MW;  D83B1595CA8378FD CRC64;

  Query Match            50.0%;  Score 42;  DB 1;  Length 653;
  Best Local Similarity  61.5%;  Pred. No. 30;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDE 13
              | || | ||:.||
Db        186 QEYGHEERRVQDE 198


RESULT 9
GVK1_HALN1
ID   GVK1_HALN1     STANDARD;      PRT;   113 AA.
```

```
AC   P24375; Q9HI26;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   GvpK protein 1.
GN   (GVPK11 OR GVPK OR VNG5021G) AND (GVPK12 OR VNG6021G).
OS   Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081), and
OS   Halobacterium sp. (strain NRC-817).
OG   Plasmid pNRC100, Plasmid pNRC200, and Plasmid pHH1.
OC   Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC   Halobacteriaceae; Halobacterium.
OX   NCBI_TaxID=64091, 148370;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRC-1; PLASMID=pNRC100;
RX   MEDLINE=91323716; PubMed=1864501;
RA   Jones J.G., Young D.C., Dassarma S.;
RT   "Structure and organization of the gas vesicle gene cluster on the
RT   Halobacterium halobium plasmid pNRC100.";
RL   Gene 102:117-122(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRC-1; PLASMID=pNRC100;
RX   MEDLINE=99063795; PubMed=9847077;
RA   Ng W.V., Ciufo S.A., Smith T.M., Bumgarner R.E., Baskin D., Faust J.,
RA   Hall B., Loretz C., Seto J., Slagel J., Hood L., DasSarma S.;
RT   "Snapshot of a large dynamic replicon in a halophilic archaeon:
RT   megaplasmid or minichromosome?";
RL   Genome Res. 8:1131-1141(1998).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRC-1; PLASMID=pNRC200;
RX   MEDLINE=20504483; PubMed=11016950;
RA   Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA   Shukla H.D., Lasky S.R., Baliga N.S., Thorsson V., Sbrogna J.,
RA   Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA   Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA   Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA   Isenbarger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA   Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA   Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.;
RT   "Genome sequence of Halobacterium species NRC-1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NRC-817; PLASMID=pHH1;
RX   MEDLINE=92065812; PubMed=1956294;
RA   Horne M., Englert C., Wimmer C., Pfeifer F.;
RT   "A DNA region of 9 kbp contains all genes necessary for gas vesicle
RT   synthesis in halophilic archaebacteria.";
RL   Mol. Microbiol. 5:1159-1174(1991).
CC   -!- FUNCTION: MAY PLAY A STRUCTURAL OR REGULATORY ROLE IN GAS
CC       VESICLE SYNTHESIS.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; M58557; AAA98188.1; -.
DR   EMBL; AF016485; AAC82801.1; -.
DR   EMBL; AE005141; AAG20718.1; -.
DR   EMBL; X55648; CAA39178.1; -.
DR   PIR; T08234; T08234.
DR   Pfam; PF05121; GvpK; 1.
KW   Gas vesicle; Plasmid; Complete proteome.
SQ   SEQUENCE   113 AA;  12695 MW;  97A469D2C1643ABF CRC64;
```

```
  Query Match           48.8%;  Score 41;  DB 1;  Length 113;
  Best Local Similarity 46.7%;  Pred. No. 7;
  Matches    7;  Conservative    4;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFE 15
              :| ||:|: : || |
Db         50 ERLGRQLQALEDELE 64


RESULT 10
6PGL_THEMA
ID   6PGL_THEMA     STANDARD;      PRT;   220 AA.
AC   Q9X0N8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN   PGL OR DEVB OR TM1154.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC       PHOSPHOGLUCONATE.
CC   -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC       phospho-D-gluconate.
CC   -!- PATHWAY: Pentose phosphate pathway; second step.
CC   -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC       ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; AE001772; AAD36230.1; -.
DR   PIR; F72289; F72289.
DR   TIGR; TM1154; -.
DR   InterPro; IPR006148; Gluc_gal_isom.
DR   InterPro; IPR005900; Phosphogluconlac.
DR   Pfam; PF01182; Glucosamine_iso; 1.
DR   TIGRFAMs; TIGR01198; pgl; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   220 AA;  25325 MW;  9B0FD07EE01E60C3 CRC64;

  Query Match           48.8%;  Score 41;  DB 1;  Length 220;
  Best Local Similarity 40.0%;  Pred. No. 14;
  Matches    6;  Conservative    6;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFE 15
              ::| ||:|  :|:|:
Db        113 EKYEREIRSATDQFD 127


RESULT 11
SNF4_YEAST
ID   SNF4_YEAST     STANDARD;      PRT;   322 AA.
AC   P12904;
DT   01-OCT-1989 (Rel. 12, Created)
DT   01-OCT-1989 (Rel. 12, Last sequence update)
```

```
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nuclear protein SNF4 (Regulatory protein CAT3).
GN   SNF4 OR CAT3 OR YGL115W.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89006284; PubMed=3049255;
RA   Schueller H.-J., Entian K.-D.;
RT   "Molecular characterization of yeast regulatory gene CAT3 necessary
RT   for glucose derepression and nuclear localization of its product.";
RL   Gene 67:247-257(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90097921; PubMed=2481228;
RA   Celenza J.L., Eng F.J., Carlson M.;
RT   "Molecular analysis of the SNF4 gene of Saccharomyces cerevisiae:
RT   evidence for physical association of the SNF4 protein with the SNF1
RT   protein kinase.";
RL   Mol. Cell. Biol. 9:5045-5054(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Lauquin G.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE OF 1-21 FROM N.A.
RA   Doi A., Doi K.;
RL   Submitted (JUN-1993) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 30-34 AND 315-322.
RX   MEDLINE=94131988; PubMed=7905477;
RA   Mitchelhill K.I., Stapleton D., Gao G., House C., Michell B.,
RA   Katsis F., Witters L.A., Kemp B.E.;
RT   "Mammalian AMP-activated protein kinase shares structural and
RT   functional homology with the catalytic domain of yeast Snf1 protein
RT   kinase.";
RL   J. Biol. Chem. 269:2361-2364(1994).
CC   -!- FUNCTION: THIS PROTEIN CAUSES EXPRESSION OF GLUCOSE-REPRESSIBLE
CC       GENES UPON GLUCOSE DEPRIVATION. IT INTERACTS AND HAS FUNCTIONAL
CC       RELATIONSHIP TO THE PROTEIN-KINASE SNF1.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC       SUBUNIT FAMILY.
CC   -!- SIMILARITY: Contains 4 CBS domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M21760; AAA34472.1; -.
DR   EMBL; M30470; AAA35061.1; -.
DR   EMBL; Z72637; CAA96823.1; -.
DR   EMBL; D16506; BAA03958.1; -.
DR   PIR; A38906; RGBYC3.
DR   SGD; S0003083; SNF4.
DR   GO; GO:0005737; C:cytoplasm; IDA.
DR   GO; GO:0005634; C:nucleus; IDA.
DR   GO; GO:0030295; F:protein kinase activator activity; IGI.
DR   GO; GO:0007031; P:peroxisome organization and biogenesis; IMP.
DR   GO; GO:0006357; P:regulation of transcription from Pol II pro. . .; IGI.
DR   InterPro; IPR000644; CBS_domain.
DR   Pfam; PF00571; CBS; 4.
DR   SMART; SM00116; CBS; 4.
KW   Carbohydrate metabolism; Transcription regulation; Nuclear protein;
KW   Repeat; CBS domain.
FT   DOMAIN       35     89       CBS 1.
FT   DOMAIN      117    175       CBS 2.
FT   DOMAIN      192    246       CBS 3.
FT   DOMAIN      259    318       CBS 4.
SQ   SEQUENCE   322 AA;  36401 MW;  51B387E346EE9561 CRC64;

  Query Match            48.8%;  Score 41;  DB 1;  Length 322;
  Best Local Similarity  61.5%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              |  | | ||:|||
Db        258 GEALMRRSDDFEG 270


RESULT 12
SNF4_KLULA
ID   SNF4_KLULA     STANDARD;      PRT;   328 AA.
AC   Q9P869;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Nuclear protein SNF4.
GN   SNF4.
OS   Kluyveromyces lactis (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX   NCBI_TaxID=28985;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Tomasini L., Ferrero I., Goffrini P.;
RT   "Molecular characterization of KlSNF4 gene.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THIS PROTEIN CAUSES EXPRESSION OF GLUCOSE-REPRESSIBLE
CC       GENES UPON GLUCOSE DEPRIVATION. IT INTERACTS AND HAS FUNCTIONAL
CC       RELATIONSHIP TO THE PROTEIN-KINASE SNF1 (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE 5'-AMP-ACTIVATED PROTEIN KINASE, GAMMA
CC       SUBUNIT FAMILY.
CC   -!- SIMILARITY: Contains 4 CBS domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AJ277480; CAB89520.1; -.
DR   InterPro; IPR000644; CBS_domain.
DR   Pfam; PF00571; CBS; 4.
DR   SMART; SM00116; CBS; 4.
KW   Carbohydrate metabolism; Transcription regulation; Nuclear protein;
KW   Repeat; CBS domain.
FT   DOMAIN       40     94       CBS 1.
FT   DOMAIN      122    180       CBS 2.
FT   DOMAIN      198    252       CBS 3.
FT   DOMAIN      271    324       CBS 4.
SQ   SEQUENCE   328 AA;  37163 MW;  DC9ED3F85E46BAD3 CRC64;

  Query Match            48.8%;  Score 41;  DB 1;  Length 328;
  Best Local Similarity  61.5%;  Pred. No. 21;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              |  | | ||:|||
Db        264 GEALMRRSDDFEG 276


RESULT 13
PSMR_ARCFU
ID   PSMR_ARCFU     STANDARD;      PRT;   398 AA.
AC   O28303;
DT   15-JUL-1998 (Rel. 36, Created)
```

```
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Proteasome-activating nucleotidase (Proteasome regulatory subunit).
GN   PAN OR AF1976.
OS   Archaeoglobus fulgidus.
OC   Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC   Archaeoglobaceae; Archaeoglobus.
OX   NCBI_TaxID=2234;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=VC-16 / DSM 4304 / ATCC 49558;
RX   MEDLINE=98049343; PubMed=9389475;
RA   Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA   Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA   Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA   Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA   Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA   Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA   Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA   Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA   Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA   Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA   Venter J.C.;
RT   "The complete genome sequence of the hyperthermophilic, sulphate-
RT   reducing archaeon Archaeoglobus fulgidus.";
RL   Nature 390:364-370(1997).
CC   -!- FUNCTION: Required for the ATP- or CTP-dependent degradation of
CC       proteins, but not small peptides, by the 20S proteasome (By
CC       similarity).
CC   -!- SUBUNIT: Homohexamer (Potential).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC   -!- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE000967; AAB89280.1; -.
DR   PIR; G69496; G69496.
DR   TIGR; AF1976; -.
DR   HAMAP; MF_00553; -; 1.
DR   InterPro; IPR005937; 26S_p45.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003959; AAA_ATPase_centr.
DR   InterPro; IPR003960; AAA_sub.
DR   Pfam; PF00004; AAA; 1.
DR   SMART; SM00382; AAA; 1.
DR   TIGRFAMs; TIGR01242; 26Sp45; 1.
DR   PROSITE; PS00674; AAA; 1.
KW   Proteasome; ATP-binding; Complete proteome.
FT   NP_BIND     182    189       ATP (POTENTIAL).
SQ   SEQUENCE   398 AA;  44964 MW;  F3293BB7D6A646B4 CRC64;

  Query Match            48.8%;  Score 41;  DB 1;  Length 398;
  Best Local Similarity   57.1%;  Pred. No. 26;
  Matches    8;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFE 15
              || ||:||:  | |
Db         43 RYEREVRRLRSEVE 56


RESULT 14
FMRA_CALPA
ID   FMRA_CALPA     STANDARD;      PRT;   334 AA.
AC   Q01133;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Antho-RFamide neuropeptides precursor.
OS   Calliactis parasitica (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Hormathiidae; Calliactis.
OX   NCBI_TaxID=6114;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91172845; PubMed=1706527;
RA   Darmer D., Schmutzler C., Diekhoff D., Grimmelikhuijzen C.J.P.;
RT   "Primary structure of the precursor for the sea anemone neuropeptide
RT   Antho-RFamide (<Glu-Gly-Arg-Phe-NH2).";
RL   Proc. Natl. Acad. Sci. U.S.A. 88:2555-2559(1991).
CC   -!- FUNCTION: NOT KNOWN BUT IT COULD ACT AS A TRANSMITTER AT
CC       NEUROMUSCULAR SYNAPSES.
CC   -!- TISSUE SPECIFICITY: NEURONS ASSOCIATED WITH SMOOTH MUSCLE FIBERS.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M59166; AAA27878.1; -.
DR   PIR; A39172; A39172.
DR   InterPro; IPR002544; FARP.
DR   Pfam; PF01581; FARP; 15.
KW   Neuropeptide; Amidation; Repeat; Signal.
FT   SIGNAL       1     26       POTENTIAL.
FT   PEPTIDE    117    120       ANTHO-RFAMIDE.
FT   PEPTIDE    126    129       ANTHO-RFAMIDE.
FT   PEPTIDE    135    138       ANTHO-RFAMIDE.
FT   PEPTIDE    143    146       ANTHO-RFAMIDE.
FT   PEPTIDE    152    155       ANTHO-RFAMIDE.
FT   PEPTIDE    161    164       ANTHO-RFAMIDE.
FT   PEPTIDE    170    173       ANTHO-RFAMIDE.
FT   PEPTIDE    179    182       ANTHO-RFAMIDE.
FT   PEPTIDE    188    191       ANTHO-RFAMIDE.
FT   PEPTIDE    197    200       ANTHO-RFAMIDE.
FT   PEPTIDE    206    209       ANTHO-RFAMIDE.
FT   PEPTIDE    215    218       ANTHO-RFAMIDE.
FT   PEPTIDE    224    227       ANTHO-RFAMIDE.
FT   PEPTIDE    234    237       ANTHO-RFAMIDE.
FT   PEPTIDE    243    246       ANTHO-RFAMIDE.
FT   PEPTIDE    253    256       ANTHO-RFAMIDE.
FT   PEPTIDE    263    266       ANTHO-RFAMIDE.
FT   PEPTIDE    272    275       ANTHO-RFAMIDE.
FT   PEPTIDE    281    284       ANTHO-RFAMIDE.
FT   MOD_RES    120    120       AMIDATION (G-121 PROVIDE AMIDE GROUP).
FT   MOD_RES    129    129       AMIDATION (G-130 PROVIDE AMIDE GROUP).
FT   MOD_RES    138    138       AMIDATION (G-139 PROVIDE AMIDE GROUP).
FT   MOD_RES    146    146       AMIDATION (G-147 PROVIDE AMIDE GROUP).
FT   MOD_RES    155    155       AMIDATION (G-156 PROVIDE AMIDE GROUP).
FT   MOD_RES    164    164       AMIDATION (G-165 PROVIDE AMIDE GROUP).
FT   MOD_RES    173    173       AMIDATION (G-174 PROVIDE AMIDE GROUP).
FT   MOD_RES    182    182       AMIDATION (G-183 PROVIDE AMIDE GROUP).
FT   MOD_RES    191    191       AMIDATION (G-192 PROVIDE AMIDE GROUP).
FT   MOD_RES    200    200       AMIDATION (G-201 PROVIDE AMIDE GROUP).
FT   MOD_RES    209    209       AMIDATION (G-210 PROVIDE AMIDE GROUP).
FT   MOD_RES    218    218       AMIDATION (G-219 PROVIDE AMIDE GROUP).
FT   MOD_RES    227    227       AMIDATION (G-228 PROVIDE AMIDE GROUP).
FT   MOD_RES    237    237       AMIDATION (G-238 PROVIDE AMIDE GROUP).
FT   MOD_RES    246    246       AMIDATION (G-247 PROVIDE AMIDE GROUP).
FT   MOD_RES    256    256       AMIDATION (G-257 PROVIDE AMIDE GROUP).
FT   MOD_RES    266    266       AMIDATION (G-267 PROVIDE AMIDE GROUP).
FT   MOD_RES    275    275       AMIDATION (G-276 PROVIDE AMIDE GROUP).
FT   MOD_RES    284    284       AMIDATION (G-285 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   334 AA;  39781 MW;  438E182C736EB583 CRC64;

  Query Match            48.2%;  Score 40.5;  DB 1;  Length 334;
  Best Local Similarity   56.2%;  Pred. No. 26;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    1;  Gaps    1;
```

```
Qy          2 RYGRELR-RMSDEFEG 16
              ||||| : |   ||:|
Db         96 RYGREFQGRFGREFQG 111


RESULT 15
FMR2_ANTEL
ID   FMR2_ANTEL     STANDARD;      PRT;   429 AA.
AC   Q16994;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Antho-RFamide neuropeptides type 2 precursor.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93054550; PubMed=1429603;
RA   Schmutzler C., Darmer D., Diekhoff D., Grimmelikhuijzen C.J.P.;
RT   "Identification of a novel type of processing sites in the precursor
RT   for the sea anemone neuropeptide Antho-RFamide (<Glu-Gly-Arg-Phe-NH2)
RT   from Anthopleura elegantissima.";
RL   J. Biol. Chem. 267:22534-22541(1992).
RN   [2]
RP   PARTIAL SEQUENCE (ANTHO-RFAMIDE).
RX   MEDLINE=87092339; PubMed=2879288;
RA   Grimmelikhuijzen C.J.P., Graff D.;
RT   "Isolation of pyroGlu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide
RT   from sea anemones.";
RL   Proc. Natl. Acad. Sci. U.S.A. 83:9817-9821(1986).
CC   -!- FUNCTION: NOT KNOWN BUT IT COULD ACT AS A TRANSMITTER AT
CC       NEUROMUSCULAR SYNAPSES.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M99170; AAA27739.1; -.
DR   Pfam; PF01581; FARP; 14.
KW   Neuropeptide; Amidation; Repeat; Signal.
FT   SIGNAL       1     22       POTENTIAL.
FT   PEPTIDE    234    237       ANTHO-RFAMIDE.
FT   PEPTIDE    242    245       ANTHO-RFAMIDE.
FT   PEPTIDE    250    253       ANTHO-RFAMIDE.
FT   PEPTIDE    258    261       ANTHO-RFAMIDE.
FT   PEPTIDE    266    269       ANTHO-RFAMIDE.
FT   PEPTIDE    274    277       ANTHO-RFAMIDE.
FT   PEPTIDE    290    293       ANTHO-RFAMIDE.
FT   PEPTIDE    298    301       ANTHO-RFAMIDE.
FT   PEPTIDE    306    309       ANTHO-RFAMIDE.
FT   PEPTIDE    322    325       ANTHO-RFAMIDE.
FT   PEPTIDE    330    333       ANTHO-RFAMIDE.
FT   PEPTIDE    343    346       ANTHO-RFAMIDE.
FT   PEPTIDE    356    359       ANTHO-RFAMIDE.
FT   PEPTIDE    369    372       ANTHO-RFAMIDE.
FT   MOD_RES    237    237       AMIDATION (G-238 PROVIDE AMIDE GROUP).
FT   MOD_RES    245    245       AMIDATION (G-246 PROVIDE AMIDE GROUP).
FT   MOD_RES    253    253       AMIDATION (G-254 PROVIDE AMIDE GROUP).
FT   MOD_RES    261    261       AMIDATION (G-262 PROVIDE AMIDE GROUP).
FT   MOD_RES    269    269       AMIDATION (G-270 PROVIDE AMIDE GROUP).
FT   MOD_RES    277    277       AMIDATION (G-278 PROVIDE AMIDE GROUP).
FT   MOD_RES    293    293       AMIDATION (G-294 PROVIDE AMIDE GROUP).
FT   MOD_RES    301    301       AMIDATION (G-302 PROVIDE AMIDE GROUP).
FT   MOD_RES    309    309       AMIDATION (G-310 PROVIDE AMIDE GROUP).
FT   MOD_RES    325    325       AMIDATION (G-326 PROVIDE AMIDE GROUP).
FT   MOD_RES    333    333       AMIDATION (G-334 PROVIDE AMIDE GROUP).
FT   MOD_RES    346    346       AMIDATION (G-347 PROVIDE AMIDE GROUP).
FT   MOD_RES    359    359       AMIDATION (G-360 PROVIDE AMIDE GROUP).
FT   MOD_RES    372    372       AMIDATION (G-373 PROVIDE AMIDE GROUP).
SQ   SEQUENCE   429 AA;  50564 MW;  7C54F5C606D537F4 CRC64;

  Query Match             48.2%;  Score 40.5;  DB 1;  Length 429;
  Best Local Similarity   56.2%;  Pred. No. 34;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          2 RYGRELR-RMSDEFEG 16
              |:|||: |   ||:|
Db        212 RFGRELQGRFGREFQG 227


Search completed: September 15, 2003, 17:23:00
Job time : 4.77143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 17.3714 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-28
Perfect score: 84
Sequence: 1 QRYGRELRRMSDEFEG 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	61	72.6	146	13	Q9I9N2	Q9i9n2 brachydanio
2	47	56.0	216	16	Q9KZC5	Q9kzc5 streptomyce
3	45	53.6	230	16	Q8XXS6	Q8xxs6 ralstonia s
4	45	53.6	567	10	Q9FV20	Q9fv20 oryza sativ
5	45	53.6	1696	10	Q8GTN3	Q8gtn3 oryza sativ
6	45	53.6	1748	16	Q989R4	Q989r4 rhizobium l
7	45	53.6	1798	2	Q8KGY8	Q8kgy8 rhizobium l
8	44	52.4	157	16	Q8ES55	Q8es55 oceanobacil
9	44	52.4	564	16	Q9RUK9	Q9ruk9 deinococcus
10	43.5	51.8	1589	2	Q8KUF5	Q8kuf5 actinosynne
11	43	51.2	463	16	Q9HUS7	Q9hus7 pseudomonas
12	43	51.2	674	10	Q93Y86	Q93y86 oryza sativ
13	43	51.2	5635	5	Q9N9N1	Q9n9n1 leishmania
14	42.5	50.6	904	2	Q9KGW3	Q9kgw3 pseudomonas
15	42.5	50.6	909	16	Q8EI34	Q8ei34 shewanella
16	42	50.0	164	17	Q8ZTI1	Q8zti1 pyrobaculum
17	42	50.0	213	17	Q8TJ31	Q8tj31 methanosarc
18	42	50.0	337	3	Q9P5V2	Q9p5v2 neurospora
19	42	50.0	375	10	Q94H58	Q94h58 oryza sativ
20	42	50.0	425	10	Q8S6J1	Q8s6j1 oryza sativ
21	42	50.0	445	16	Q97I40	Q97i40 clostridium
22	42	50.0	561	10	Q8LM44	Q8lm44 oryza sativ
23	42	50.0	581	16	Q8F646	Q8f646 leptospira
24	42	50.0	606	11	Q8C850	Q8c850 mus musculu
25	42	50.0	606	11	Q8C6N1	Q8c6n1 mus musculu
26	42	50.0	641	10	Q8GTN7	Q8gtn7 oryza sativ
27	42	50.0	677	10	Q9AYG6	Q9ayg6 oryza sativ
28	42	50.0	804	16	Q8EW78	Q8ew78 mycoplasma
29	42	50.0	827	10	Q8GTN2	Q8gtn2 oryza sativ
30	42	50.0	829	10	Q9XEM4	Q9xem4 oryza sativ
31	42	50.0	829	10	Q8GVG4	Q8gvg4 oryza sativ
32	42	50.0	1032	10	Q9LGR7	Q9lgr7 oryza sativ
33	42	50.0	1032	10	Q8LHG3	Q8lhg3 oryza sativ
34	42	50.0	1032	10	Q9LGR6	Q9lgr6 oryza sativ
35	42	50.0	1032	10	Q9SSW6	Q9ssw6 oryza sativ
36	42	50.0	1032	10	Q8RYW0	Q8ryw0 oryza sativ
37	42	50.0	1032	10	Q8GTN5	Q8gtn5 oryza sativ
38	42	50.0	1099	10	Q8W3I5	Q8w3i5 oryza sativ
39	42	50.0	1112	10	Q8LSS6	Q8lss6 oryza sativ
40	42	50.0	1126	5	Q9VSM2	Q9vsm2 drosophila
41	42	50.0	1416	10	Q8S609	Q8s609 oryza sativ
42	42	50.0	1429	10	Q94HU1	Q94hu1 oryza sativ
43	42	50.0	1583	10	Q9S7A7	Q9s7a7 oryza sativ
44	42	50.0	1653	5	Q9VSM1	Q9vsm1 drosophila
45	42	50.0	1709	5	Q9VSM0	Q9vsm0 drosophila

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 72.6%; Score 61; DB 13; Length 146;
Best Local Similarity 66.7%; Pred. No. 0.028;
Matches 10; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEFE 15
              ::||::||||||||:
Db         93 KKYGQQLRRMSDEFD 107
```

RESULT 2
Q9KZC5
ID Q9KZC5 PRELIMINARY; PRT; 216 AA.
AC Q9KZC5;

```
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein SCO6964.
GN   SCO6964 OR SC6F7.17C.
OS   Streptomyces coelicolor.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1902;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Saunders D.C., Harris D.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RA   Thomson N.R., Parkhill J., Barrell B.G., Rajandream M.A.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2);
RX   MEDLINE=97000351; PubMed=8843436;
RA   Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA   Kinashi H., Hopwood D.A.;
RT   "A set of ordered cosmids and a detailed genetic and physical map for
RT   the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL   Mol. Microbiol. 21:77-96(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=A3(2) / M145;
RX   MEDLINE=21996410; PubMed=12000953;
RA   Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA   Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA   Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA   Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA   Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neil S.,
RA   Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA   Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA   Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA   Hopwood D.A.;
RT   "Complete genome sequence of the model actinomycete Streptomyces
RT   coelicolor A3(2).";
RL   Nature 417:141-147(2002).
DR   EMBL; AL939129; CAB89025.1; -.
DR   InterPro; IPR003265; Endo_3c.
DR   Pfam; PF00730; HhH-GPD; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   216 AA;  23810 MW;  DDD7C717D60F7AA7 CRC64;

  Query Match             56.0%;  Score 47;  DB 16;  Length 216;
  Best Local Similarity   50.0%;  Pred. No. 8.7;
  Matches    8;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              :|:| :|||: || :|
Db        109 ERWGGDLRRLRDEADG 124


RESULT 3
Q8XXS6
ID   Q8XXS6      PRELIMINARY;      PRT;   230 AA.
AC   Q8XXS6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Probable ATP-binding ABC transporter protein.
GN   RSC2037 OR RS03602.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Ralstoniaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646067; CAD15739.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
KW   Complete proteome.
SQ   SEQUENCE   230 AA;  25231 MW;  7C3FDA1E7A19A2A4 CRC64;

  Query Match             53.6%;  Score 45;  DB 16;  Length 230;
  Best Local Similarity   64.3%;  Pred. No. 20;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              |  || |||: |||
Db        168 QEIGRTLRRLVDEF 181


RESULT 4
Q9FV20
ID   Q9FV20      PRELIMINARY;      PRT;   567 AA.
AC   Q9FV20;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Polyprotein.
OS   Oryza sativa (indica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39946;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Zhaiyeqing 8; TRANSPOSON=unspecified; TISSUE=Shoot;
RA   Yu F., Zhang A., Zhang F., Chen S.;
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF259976; AAG15480.1; -.
DR   Gramene; Q9FV20; -.
DR   InterPro; IPR002156; RNaseH.
DR   InterPro; IPR000477; RVTse.
DR   Pfam; PF00075; rnaseH; 1.
DR   Pfam; PF00078; rvt; 1.
KW   RNA-directed DNA polymerase; Transferase.
SQ   SEQUENCE   567 AA;  64419 MW;  BD4F47AB33A685BE CRC64;

  Query Match             53.6%;  Score 45;  DB 10;  Length 567;
  Best Local Similarity   50.0%;  Pred. No. 52;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFEG 16
              | :|:|:: |:|||
Db        547 YRQEVRKLEDKFEG 560


RESULT 5
Q8GTN3
ID   Q8GTN3      PRELIMINARY;      PRT;  1696 AA.
AC   Q8GTN3;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Putative GAG-POL precursor.
```

```
GN   OSJNBA0041L14.29.
OS   Oryza sativa (japonica cultivar-group).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=39947;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA   Overton II L.L., Tsitrin T., Kim M.M., Bera J.J., Jin S.S.,
RA   Fadrosh D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA   Vanaken S.S., Riedmuller S.B., Utterback T.T., Feldblyum T.V.,
RA   Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA   White O., Salzberg S.L., Fraser C.M.;
RT   "Oryza sativa chromosome 10 BAC OSJNBa0041L14 genomic sequence.";
RL   Submitted (SEP-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC099042; AAN08247.1; -.
SQ   SEQUENCE   1696 AA;  190501 MW;  516B35CFDFA2F068 CRC64;

  Query Match           53.6%;  Score 45;  DB 10;  Length 1696;
  Best Local Similarity  50.0%;  Pred. No. 1.7e+02;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFEG 16
              | :|:|:: |:|||
Db       1341 YRQEVRKLEDKFEG 1354


RESULT 6
Q989R4
ID   Q989R4      PRELIMINARY;      PRT;  1748 AA.
AC   Q989R4;
DT   01-OCT-2001 (TrEMBLrel. 18, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein mlr6316.
GN   MLR6316.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
DR   EMBL; AP003008; BAB52630.1; -.
DR   InterPro; IPR003653; SUMO_protease.
DR   Pfam; PF02902; Peptidase_C48; 1.
DR   PROSITE; PS50600; ULP_PROTEASE; 1.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   1748 AA;  190454 MW;  2A9CE7B92C354341 CRC64;

  Query Match           53.6%;  Score 45;  DB 16;  Length 1748;
  Best Local Similarity  69.2%;  Pred. No. 1.7e+02;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFE 15
              ||| ||| |:| |
Db        562 YGRLLRRFSEELE 574


RESULT 7
Q8KGY8
ID   Q8KGY8      PRELIMINARY;      PRT;  1798 AA.
AC   Q8KGY8;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   MSI059.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R7A;
RX   MEDLINE=21999272; PubMed=12003951;
RA   Sullivan J.T., Trzebiatowski J.R., Cruickshank R.W., Gouzy J.,
RA   Brown S.D., Elliot R.M., Fleetwood D.J., McCallum N.G., Rossbach U.,
RA   Stuart G.S., Weaver J.E., Webby R.J., de Bruijn F.J., Ronson C.W.;
RT   "Comparative sequence analysis of the symbiosis island of
RT   Mesorhizobium loti strain R7A.";
RL   J. Bacteriol. 184:3086-3095(2002).
DR   EMBL; AL672112; CAD31464.1; -.
DR   InterPro; IPR003653; SUMO_protease.
DR   Pfam; PF02902; Peptidase_C48; 1.
DR   PROSITE; PS50600; ULP_PROTEASE; 1.
KW   Hypothetical protein.
SQ   SEQUENCE   1798 AA;  196214 MW;  A750A49B8C0D581B CRC64;

  Query Match           53.6%;  Score 45;  DB 2;  Length 1798;
  Best Local Similarity  69.2%;  Pred. No. 1.8e+02;
  Matches    9;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFE 15
              ||| ||| |:| |
Db        569 YGRLLRRFSEELE 581


RESULT 8
Q8ES55
ID   Q8ES55      PRELIMINARY;      PRT;   157 AA.
AC   Q8ES55;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical conserved protein.
GN   OB0788.
OS   Oceanobacillus iheyensis.
OC   Bacteria; Firmicutes; Bacillales; Oceanobacillus.
OX   NCBI_TaxID=182710;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HTE831 / DSM 14371 / JCM 11309;
RX   MEDLINE=22220767; PubMed=12235376;
RA   Takami H., Takaki Y., Uchiyama I.;
RT   "Genome sequence of Oceanobacillus iheyensis isolated from the Iheya
RT   Ridge and its unexpected adaptive capabilities to extreme
RT   environments.";
RL   Nucleic Acids Res. 30:3927-3935(2002).
DR   EMBL; AP004595; BAC12744.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   157 AA;  17809 MW;  128229251596C18C CRC64;

  Query Match           52.4%;  Score 44;  DB 16;  Length 157;
  Best Local Similarity  50.0%;  Pred. No. 19;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFEG 16
              : |||:|: :|| |
Db        126 FSRELKRIGEEFSG 139


RESULT 9
Q9RUK9
ID   Q9RUK9      PRELIMINARY;      PRT;   564 AA.
```

```
AC   Q9RUK9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycosyl hydrolase, family 13.
GN   DR1375.
OS   Deinococcus radiodurans.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001983; AAF10944.1; -.
DR   HSSP; P21332; 1UOK.
DR   TIGR; DR1375; -.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR006589; Alp_amyl_cat_sub.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   SMART; SM00642; Aamy; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   564 AA;  63667 MW;  B8F50B9B0DFC8D51 CRC64;

  Query Match             52.4%;  Score 44;  DB 16;  Length 564;
  Best Local Similarity   61.5%;  Pred. No. 76;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFE 15
              | ||:||: |||:
Db        283 YVREMRRVIDEFD 295


RESULT 10
Q8KUF5
ID   Q8KUF5      PRELIMINARY;      PRT;  1589 AA.
AC   Q8KUF5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Polyketide synthase.
GN   ASMC.
OS   Actinosynnema pretiosum (subsp. auranticum).
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Pseudonocardineae; Actinosynnemataceae; Actinosynnema.
OX   NCBI_TaxID=42198;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 31565;
RX   MEDLINE=22056096; PubMed=12060743;
RA   Yu T.W., Bai L., Clade D., Hoffmann D., Toelzer S., Trinh K.Q., Xu J.,
RA   Moss S.J., Leistner E., Floss H.G.;
RT   "The biosynthetic gene cluster of the maytansinoid antitumor agent
RT   ansamitocin from Actinosynnema pretiosum.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:7968-7973(2002).
DR   EMBL; AF453501; AAM54077.1; -.
DR   InterPro; IPR001227; Ac_transferase.
DR   InterPro; IPR000794; Ketoacyl-synt.
DR   InterPro; IPR006162; Ppantne_attach.
DR   InterPro; IPR006163; Pp_bind.
DR   Pfam; PF00698; Acyl_transf; 1.
DR   Pfam; PF00109; ketoacyl-synt; 1.
DR   Pfam; PF02801; ketoacyl-synt_C; 1.
DR   Pfam; PF00550; pp-binding; 1.
DR   PROSITE; PS50075; ACP_DOMAIN; 1.
DR   PROSITE; PS00606; B_KETOACYL_SYNTHASE; 1.
DR   PROSITE; PS00012; PHOSPHOPANTETHEINE; 1.
KW   Phosphopantetheine.
SQ   SEQUENCE   1589 AA;  164712 MW;  5067CF9A772A525F CRC64;

  Query Match             51.8%;  Score 43.5;  DB 2;  Length 1589;
  Best Local Similarity   56.2%;  Pred. No. 2.7e+02;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 QRYGRELRRMSDEFEG 16
              | ||   ||:::||||
Db        161 QNYGNR-RRVAEEFEG 175


RESULT 11
Q9HUS7
ID   Q9HUS7      PRELIMINARY;      PRT;  463 AA.
AC   Q9HUS7;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Probable two-component sensor.
GN   PA4886.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
CC   -!- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR   EMBL; AE004901; AAG08271.1; -.
DR   HSSP; P02933; 1BXD.
DR   InterPro; IPR003594; ATPbind_ATPase.
DR   InterPro; IPR004358; Bact_sens_pr_C.
DR   InterPro; IPR003660; HAMP.
DR   InterPro; IPR003661; His_kinA.
DR   InterPro; IPR005467; His_kinase.
DR   InterPro; IPR000005; HTHAraC.
DR   InterPro; IPR006290; Metal_his_kin.
DR   Pfam; PF00672; HAMP; 1.
DR   Pfam; PF02518; HATPase_c; 1.
DR   Pfam; PF00512; HisKA; 1.
DR   PRINTS; PR00344; BCTRLSENSOR.
DR   SMART; SM00304; HAMP; 1.
DR   SMART; SM00387; HATPase_c; 1.
DR   SMART; SM00388; HisKA; 1.
DR   TIGRFAMs; TIGR01386; cztS_silS_copS; 1.
DR   PROSITE; PS50109; HIS_KIN; 1.
DR   PROSITE; PS00041; HTH_ARAC_FAMILY_1; 1.
KW   Kinase; Phosphorylation; Sensory transduction; Transferase;
KW   Complete proteome.
SQ   SEQUENCE   463 AA;  50946 MW;  97FF19C2CF38006C CRC64;

  Query Match             51.2%;  Score 43;  DB 16;  Length 463;
  Best Local Similarity   61.5%;  Pred. No. 90;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFEG 16
              | |||| :: |||
Db        321 GEELRRQAEYFEG 333
```

```
RESULT 12
Q93Y86
ID   Q93Y86      PRELIMINARY;      PRT;   674 AA.
AC   Q93Y86;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Putative gag-pol.
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare;
RA   Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,
RA   Brenner M., Burgess S., Hance M., Shvartsbeyn M., Tsitrin T.,
RA   Riggs F., Hsiao J., Zismann V., Blunt S., Pai G., VanAken S.E.,
RA   Utterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,
RA   White O., Fraser C.M.;
RT   "Oryza sativa chromosome 3 BAC OSJNBa0026A15 genomic sequence.";
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC084404; AAK50597.1; -.
DR   Gramene; Q93Y86; -.
DR   InterPro; IPR002156; RNaseH.
DR   InterPro; IPR001584; Rve.
DR   Pfam; PF00075; rnaseH; 1.
DR   Pfam; PF00665; rve; 1.
SQ   SEQUENCE   674 AA;  75473 MW;  9FA929B1426E725E CRC64;

  Query Match             51.2%;  Score 43;  DB 10;  Length 674;
  Best Local Similarity   50.0%;  Pred. No. 1.3e+02;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFEG 16
              : | ||:|||   |:|
Db        258 EAYVREVRRMERNFDG 273


RESULT 13
Q9N9N1
ID   Q9N9N1      PRELIMINARY;      PRT;  5635 AA.
AC   Q9N9N1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Dynein heavy chain, cytosolic.
GN   L3302.02.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RA   Hilbert H., Wedler H., Wedler E., Duesterhoeft A., Ivens A.C.,
RA   Quail M., Rajandream M.A., Barrell B.G.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL359781; CAB95305.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   InterPro; IPR000169; SHprot_acsite.
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 4.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW   ATP-binding.
SQ   SEQUENCE   5635 AA;  620050 MW;  64A9EB81A9B14641 CRC64;

  Query Match             51.2%;  Score 43;  DB 5;  Length 5635;
  Best Local Similarity   72.7%;  Pred. No. 1.3e+03;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMS 11
              ||:||:||| |
Db       1535 QRFGRDLRRWS 1545


RESULT 14
Q9KGW3
ID   Q9KGW3      PRELIMINARY;      PRT;   904 AA.
AC   Q9KGW3;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase I subunit G.
GN   NUOG.
OS   Pseudomonas fluorescens.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=294;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WCS365;
RA   Camacho Carvajal M.M., Lugtenberg B.J.J., Bloemberg G.V.;
RT   "Characterization of NADH dehydrogenases of Pseudomonas fluorescens
RT   WCS365 and their role in competitive root colonisation.";
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF281148; AAF97803.1; -.
DR   InterPro; IPR000283; Complex1_75K.
DR   InterPro; IPR001041; Ferredoxin.
DR   InterPro; IPR006656; Molybdopterin.
DR   InterPro; IPR006963; Molybdop_Fe4S4.
DR   Pfam; PF00111; fer2; 1.
DR   Pfam; PF00384; molybdopterin; 1.
DR   Pfam; PF04879; Molybdop_Fe4S4; 1.
DR   PROSITE; PS00641; COMPLEX1_75K_1; 1.
DR   PROSITE; PS00642; COMPLEX1_75K_2; 1.
DR   PROSITE; PS00643; COMPLEX1_75K_3; 1.
KW   Iron; Iron-sulfur.
SQ   SEQUENCE   904 AA;  98157 MW;  C25E86C6D4DFA457 CRC64;

  Query Match             50.6%;  Score 42.5;  DB 2;  Length 904;
  Best Local Similarity   56.2%;  Pred. No. 2.2e+02;
  Matches    9;  Conservative    3;  Mismatches    3;  Indels    1;  Gaps    1;

Qy          1 QRYGRELRRMSDEFEG 16
              :||| |||| : | |
Db        241 ERYG-ELRRIENRFNG 255


RESULT 15
Q8EI34
ID   Q8EI34      PRELIMINARY;      PRT;   909 AA.
AC   Q8EI34;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   NADH dehydrogenase I, G subunit.
GN   NUOG OR SO1016.
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Alteromonadaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MR-1;
```

RX MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Tsapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Nealson K.H., Fraser C.M.;
RT "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis.";
RL Nat. Biotechnol. 20:1118-1123(2002).
DR EMBL; AE015546; AAN54089.1; -.
DR TIGR; SO1016; -.
KW Complete proteome.
SQ SEQUENCE 909 AA; 100079 MW; F71859F385BDCAC8 CRC64;

Query Match 50.6%; Score 42.5; DB 16; Length 909;
Best Local Similarity 56.2%; Pred. No. 2.2e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 1; Gaps 1;

```
Qy          1 QRYGRELRRMSDEFEG 16
              :||| |||| : | |
Db        241 ERYG-ELRRIENRFHG 255
```

Search completed: September 15, 2003, 17:25:48
Job time : 19.3714 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:01 ; Search time 22.6286 Seconds
(without alignments)
112.231 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	16	20	AAY05421	Human BAD BH3 doma
2	83	100.0	16	21	AAB37029	Bcl2 polypeptide B
3	83	100.0	20	23	ABG78499	Mutant Bcl2 compet
4	83	100.0	20	23	AAU78626	Human Bad peptide
5	83	100.0	21	23	ABG78500	Mutant Bcl2 compet
6	83	100.0	21	23	AAU78630	Human Bad peptide
7	83	100.0	22	23	AAU78629	Human Bad peptide
8	83	100.0	23	23	AAU78628	Human Bad peptide
9	83	100.0	24	23	ABG78482	Human Bcl2 fluores
10	83	100.0	24	23	AAU78605	Human Bad peptide
11	83	100.0	24	23	AAU78627	Human Bad peptide
12	83	100.0	25	23	ABP56161	PTPC-interacting T
13	83	100.0	25	23	ABG78481	Human Bcl2 fluores
14	83	100.0	25	23	ABG78484	Mutant Bcl2 compet
15	83	100.0	25	23	ABG78485	Mutant Bcl2 compet
16	83	100.0	25	23	ABG78486	Mutant Bcl2 compet
17	83	100.0	25	23	ABG78487	Mutant Bcl2 compet
18	83	100.0	25	23	ABG78488	Mutant Bcl2 compet
19	83	100.0	25	23	ABG78489	Mutant Bcl2 compet
20	83	100.0	25	23	ABG78490	Mutant Bcl2 compet
21	83	100.0	25	23	ABG78491	Mutant Bcl2 compet
22	83	100.0	25	23	ABG78492	Mutant Bcl2 compet
23	83	100.0	25	23	ABG78493	Mutant Bcl2 compet
24	83	100.0	25	23	ABG78494	Mutant Bcl2 compet
25	83	100.0	25	23	ABG78495	Mutant Bcl2 compet
26	83	100.0	25	23	AAU78601	Human Bad peptide
27	83	100.0	25	23	AAU78610	Human Bad peptide
28	83	100.0	25	23	AAU78611	Human Bad peptide
29	83	100.0	25	23	AAU78612	Human Bad peptide
30	83	100.0	25	23	AAU78613	Human Bad peptide
31	83	100.0	25	23	AAU78614	Human Bad peptide
32	83	100.0	25	23	AAU78615	Human Bad peptide
33	83	100.0	25	23	AAU78616	Human Bad peptide
34	83	100.0	25	23	AAU78617	Human Bad peptide
35	83	100.0	25	23	AAU78618	Human Bad peptide
36	83	100.0	25	23	AAU78619	Human Bad peptide
37	83	100.0	25	23	AAU78620	Human Bad peptide
38	83	100.0	25	23	AAU78621	Human Bad peptide
39	83	100.0	25	23	AAU78622	Human Bad peptide
40	83	100.0	26	21	AAY96321	Mammalian Bad Bcl-
41	83	100.0	26	22	AAB70371	BAD BH3 consensus
42	83	100.0	166	18	AAW32476	BBC6 protein for r
43	83	100.0	168	19	AAW55779	Human Bcl-xL/Bcl-2
44	83	100.0	168	21	AAB13512	Human cell prolife
45	83	100.0	168	22	AAB70368	Human BAD mutant a

ALIGNMENTS

RESULT 1
AAY05421
ID AAY05421 standard; peptide; 16 AA.
XX
AC AAY05421;
XX
DT 02-JUL-1999 (first entry)
XX
DE Human BAD BH3 domain.
XX
KW BH3 domain; cell death agonist; bcl homology domain; BCL-2 family;
KW apoptosis promoter; cancer cell; virus infected cell; inflammation;
KW autoantibody producing cell; cancer; lymphoproliferative condition;
KW arthritis; autoimmune disease; therapy.
XX
OS Homo sapiens.
XX
PN WO9916787-A1.
XX
PD 08-APR-1999.
XX
PF 22-SEP-1998; 98WO-US19765.
XX
PR 07-OCT-1997; 97US-0946039.
PR 26-SEP-1997; 97US-0060133.
XX
PA (UNIW) UNIV WASHINGTON.
XX
PI Korsmeyer SJ;
XX
DR WPI; 1999-255058/21.
XX

PT Bcl homology domain 3 polypeptide
XX
PS Example 1; Fig 4; 104pp; English.
XX
CC This sequence represents the BH3 domain of human BAD.
CC The invention relates to a bcl homology domain 3 (BH3 domain),
CC derived from a proapoptotic member of the BCL-2 family. The
CC BH3 polypeptide can be used in a method for promoting apoptosis in a
CC target cell, especially where the cell is a cancer cell a virus infected
CC cell or an autoantibody producing cell. The BH3 polypeptide can be used
CC in therapeutic compositions for treating disease including cancer, other
CC lymphoproliferative conditions, arthritis, inflammation, and autoimmune
CC diseases, which may result from the down regulation of cell death
CC regulation.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 83; DB 20; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 1 QRYGRELRRMSDEFVD 16

RESULT 2
AAB37029
ID AAB37029 standard; peptide; 16 AA.
XX
AC AAB37029;
XX
DT 28-FEB-2001 (first entry)
XX
DE Bcl2 polypeptide BH3 domain peptide #29.
XX
KW Cytostatic; neuroprotective; anti-HIV; virucide; cerebroprotective;
KW cardiant; Bcl-2 superfamily; BH3 domain; cell death agonist; Bad;
KW apoptosis modulation; B cell lymphoma/leukemia 2; cancer; prostate;
KW colorectal; gastric; non-small lung; renal; thyroid; neuroblastoma;
KW melanoma; lymphocytic leukemia; neurodegenerative disorder; AIDS;
KW stroke; myocardial infarction.
XX
OS Homo sapiens.
XX
PN WO200059526-A1.
XX
PD 12-OCT-2000.
XX
PF 06-APR-2000; 2000WO-US09352.
XX
PR 07-APR-1999; 99US-0128202.
XX
PA (UYJE-) UNIV JEFFERSON THOMAS.
XX
PI Huang Z, Wang J, Zhang Z, Shan S, Lu Z;
XX
DR WPI; 2000-679325/66.
XX
PT New peptide conjugates for modulating apoptosis or for inhibiting B
PT cell lymphoma/leukemia 2 (Bcl-2) function, especially useful for
PT treating neurodegenerative disorders, stroke, or cancer -
XX
PS Claim 18; Page 18; 74pp; English.
XX
CC The invention relates to a peptide conjugate having the formula:
CC (R-X)n-peptide where n = 1-10; X = C=O, when the R-X group is attached
CC to the N-terminus of the peptide, or a side chain of the peptide where
CC the functional group of the side chain is NH2 or OH; or X = O or NH,
CC when the R-X group is attached to the C-terminus of the peptide, or a
CC side chain of the peptide, where the side chain functional group is COOH
CC or CONH2; and R = 2-18C alkyl or alkoxy, 2-14C alkylenyl containing one
CC or two double bonds, cyclobutyl, cyclopentyl, cyclohexyl optionally
CC monosubstituted with a 1-5C straight or branched chain alkyl group,
CC phenyl optionally monosubstituted with a 1-5C straight or branched chain
CC alkyl group, or benzyl. The peptides AAB37001-B37058 represent examples
CC of the peptide portion of the conjugate. The peptides represent analogues
CC of a Bcl-2 superfamily polypeptide corresponding to amino acids 72-97 of
CC the BH3 domain of the cell death agonist Bad. The peptide conjugate is
CC useful for modulating apoptosis in the cells of a subject, or for
CC reversing B cell lymphoma/leukemia 2 (Bcl-2)-mediated blockage of
CC apoptosis in cancer cells. It is also useful for inhibiting Bcl-2
CC function. In particular, the peptide conjugate is useful for treating a
CC subject afflicted with a cancer characterized by cancer cells that
CC express Bcl-2. The cancer includes prostate, colorectal, gastric,
CC non-small lung, renal or thyroid cancers, neuroblastoma, melanoma, or
CC acute or chronic lymphocytic and non-lymphocytic leukemia. The peptide
CC conjugate is also useful for treating disorders characterized by
CC increased apoptosis, e.g. neurodegenerative disorders, acquired
CC immunodeficiency syndrome (AIDS), stroke or myocardial infarction.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 83; DB 21; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 1 QRYGRELRRMSDEFVD 16

RESULT 3
ABG78499
ID ABG78499 standard; Peptide; 20 AA.
XX
AC ABG78499;
XX
DT 15-NOV-2002 (first entry)
XX
DE Mutant Bcl2 competitive binding assay peptide #16.
XX
KW Human; Bcl2; BclX1; programmed cell death; apoptosis; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200240530-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US45693.
XX
PR 20-NOV-2000; 2000US-0716395.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Petros AM, Yoon H, Nettesheim DG;
XX
DR WPI; 2002-490141/52.
XX
PT New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT useful in biological assays to identify substances that block the
PT ability of Bcl-2 to inhibit programmed cell death or apoptosis -
XX
PS Example 2; Page 17; 36pp; English.
XX
CC This invention relates to a novel mutant protein which is derived from
CC a wild type human Bcl-2 protein. The mutant is created by replacing a
CC sequence of amino acid residues comprising a flexible loop from the wild
CC type Bcl-2 protein with an amino acid sequence comprising at least two
CC acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC shown in the specification. The invention also comprises an assay for
CC identifying substances that bind to the Bcl-2 protein. The protein
CC sequences of the invention are useful in biological assays to identify
CC substances that block the ability of Bcl-2 to inhibit programmed cell

```
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   20 AA;

  Query Match           100.0%;  Score 83;  DB 23;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          1 QRYGRELRRMSDEFVD 16


RESULT 4
AAU78626
ID   AAU78626 standard; Peptide; 20 AA.
XX
AC   AAU78626;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human Bad peptide #26 which binds to a member of the Bcl-2 family.
XX
KW   Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW   ischemic injury; suppressor; BH3 domain.
XX
OS   Homo sapiens.
XX
PN   WO200220568-A2.
XX
PD   14-MAR-2002.
XX
PF   04-SEP-2001; 2001WO-US27410.
XX
PR   06-SEP-2000; 2000US-0656399.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Joseph MK,  Olejniczak ET,  Petros AM;
PI   Nettesheim DG,  Swift KM,  Matayoshi E,  Zhang H;
XX
DR   WPI; 2002-292254/33.
XX
PT   New derivatives of Bad peptide, useful for identifying compounds that
PT   bind to Bcl-2 proteins, potential agents for treating cancer and
PT   degenerative disease  -
XX
PS   Example 1; Page 14; 31pp; English.
XX
CC   The present invention relates to new peptides that are derived from a
CC   wild-type human Bad peptide and are able to bind to a member of the
CC   Bcl-2 protein family. The peptides are useful, when labelled, in
CC   competitive/displacement assays for identifying substances that bind to
CC   members of the Bcl-2 family and may induce or suppress apoptosis so are
CC   potentially useful for treating cancer (inducers) or degenerative
CC   diseases or ischemic injury (suppressors). The peptides of the invention
CC   have high helix propensity, maintain the contacts of the wild-type Bad
CC   peptide and, compared with the Bad peptide, may have better physical
CC   properties, particularly solubility. The present sequence represents one
CC   of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC   from the BH3 domain of the human wild-type Bad peptide.
XX
SQ   Sequence   20 AA;

  Query Match           100.0%;  Score 83;  DB 23;  Length 20;
  Best Local Similarity  100.0%;  Pred. No. 1.7e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          1 QRYGRELRRMSDEFVD 16


RESULT 5
ABG78500
ID   ABG78500 standard; Peptide; 21 AA.
XX
AC   ABG78500;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mutant Bcl2 competitive binding assay peptide #17.
XX
KW   Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200240530-A2.
XX
PD   23-MAY-2002.
XX
PF   15-NOV-2001; 2001WO-US45693.
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 17; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   21 AA;

  Query Match           100.0%;  Score 83;  DB 23;  Length 21;
  Best Local Similarity  100.0%;  Pred. No. 1.8e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 6
AAU78630
ID   AAU78630 standard; Peptide; 21 AA.
XX
AC   AAU78630;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human Bad peptide #30 which binds to a member of the Bcl-2 family.
XX
KW   Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW   ischemic injury; suppressor; BH3 domain.
XX
OS   Homo sapiens.
```

XX
PN WO200220568-A2.
XX
PD 14-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US27410.
XX
PR 06-SEP-2000; 2000US-0656399.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Joseph MK, Olejniczak ET, Petros AM;
PI Nettesheim DG, Swift KM, Matayoshi E, Zhang H;
XX
DR WPI; 2002-292254/33.
XX
PT New derivatives of Bad peptide, useful for identifying compounds that
PT bind to Bcl-2 proteins, potential agents for treating cancer and
PT degenerative disease -
XX
PS Claim 15; Page 18; 31pp; English.
XX
CC The present invention relates to new peptides that are derived from a
CC wild-type human Bad peptide and are able to bind to a member of the
CC Bcl-2 protein family. The peptides are useful, when labelled, in
CC competitive/displacement assays for identifying substances that bind to
CC members of the Bcl-2 family and may induce or suppress apoptosis so are
CC potentially useful for treating cancer (inducers) or degenerative
CC diseases or ischemic injury (suppressors). The peptides of the invention
CC have high helix propensity, maintain the contacts of the wild-type Bad
CC peptide and, compared with the Bad peptide, may have better physical
CC properties, particularly solubility. The present sequence represents one
CC of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC from the BH3 domain of the human wild-type Bad peptide.
XX
SQ Sequence 21 AA;

Query Match 100.0%; Score 83; DB 23; Length 21;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFVD 21

RESULT 7
AAU78629
ID AAU78629 standard; Peptide; 22 AA.
XX
AC AAU78629;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Bad peptide #29 which binds to a member of the Bcl-2 family.
XX
KW Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW ischemic injury; suppressor; BH3 domain.
XX
OS Homo sapiens.
XX
PN WO200220568-A2.
XX
PD 14-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US27410.
XX
PR 06-SEP-2000; 2000US-0656399.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Joseph MK, Olejniczak ET, Petros AM;
PI Nettesheim DG, Swift KM, Matayoshi E, Zhang H;
XX
DR WPI; 2002-292254/33.
XX
PT New derivatives of Bad peptide, useful for identifying compounds that
PT bind to Bcl-2 proteins, potential agents for treating cancer and
PT degenerative disease -
XX
PS Claim 15; Page 18; 31pp; English.
XX
CC The present invention relates to new peptides that are derived from a
CC wild-type human Bad peptide and are able to bind to a member of the
CC Bcl-2 protein family. The peptides are useful, when labelled, in
CC competitive/displacement assays for identifying substances that bind to
CC members of the Bcl-2 family and may induce or suppress apoptosis so are
CC potentially useful for treating cancer (inducers) or degenerative
CC diseases or ischemic injury (suppressors). The peptides of the invention
CC have high helix propensity, maintain the contacts of the wild-type Bad
CC peptide and, compared with the Bad peptide, may have better physical
CC properties, particularly solubility. The present sequence represents one
CC of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC from the BH3 domain of the human wild-type Bad peptide.
XX
SQ Sequence 22 AA;

Query Match 100.0%; Score 83; DB 23; Length 22;
Best Local Similarity 100.0%; Pred. No. 1.9e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFVD 21

RESULT 8
AAU78628
ID AAU78628 standard; Peptide; 23 AA.
XX
AC AAU78628;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Bad peptide #28 which binds to a member of the Bcl-2 family.
XX
KW Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW ischemic injury; suppressor; BH3 domain.
XX
OS Homo sapiens.
XX
PN WO200220568-A2.
XX
PD 14-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US27410.
XX
PR 06-SEP-2000; 2000US-0656399.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Joseph MK, Olejniczak ET, Petros AM;
PI Nettesheim DG, Swift KM, Matayoshi E, Zhang H;
XX
DR WPI; 2002-292254/33.
XX
PT New derivatives of Bad peptide, useful for identifying compounds that
PT bind to Bcl-2 proteins, potential agents for treating cancer and
PT degenerative disease -
XX
PS Claim 15; Page 18; 31pp; English.
XX
CC The present invention relates to new peptides that are derived from a
CC wild-type human Bad peptide and are able to bind to a member of the
CC Bcl-2 protein family. The peptides are useful, when labelled, in
CC competitive/displacement assays for identifying substances that bind to

```
CC   members of the Bcl-2 family and may induce or suppress apoptosis so are
CC   potentially useful for treating cancer (inducers) or degenerative
CC   diseases or ischemic injury (suppressors). The peptides of the invention
CC   have high helix propensity, maintain the contacts of the wild-type Bad
CC   peptide and, compared with the Bad peptide, may have better physical
CC   properties, particularly solubility. The present sequence represents one
CC   of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC   from the BH3 domain of the human wild-type Bad peptide.
XX
SQ   Sequence   23 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.9e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 9
ABG78482
ID   ABG78482 standard; Peptide; 24 AA.
XX
AC   ABG78482;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Human Bcl2 fluorescein labelled peptide #2.
XX
KW   Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200240530-A2.
XX
PD   23-MAY-2002.
XX
PF   15-NOV-2001; 2001WO-US45693.
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 15; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   24 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 10
AAU78605
ID   AAU78605 standard; Peptide; 24 AA.
XX
AC   AAU78605;
XX
DT   18-JUN-2002  (first entry)
XX
DE   Human Bad peptide #5 which binds to a member of the Bcl-2 family.
XX
KW   Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW   ischemic injury; suppressor; BH3 domain.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "Optionally labelled with 5-FAM,
FT                          (6-carboxy-fluorescein)"
XX
PN   WO200220568-A2.
XX
PD   14-MAR-2002.
XX
PF   04-SEP-2001; 2001WO-US27410.
XX
PR   06-SEP-2000; 2000US-0656399.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Meadows RP,  Joseph MK,  Olejniczak ET,  Petros AM;
PI   Nettesheim DG,  Swift KM,  Matayoshi E,  Zhang H;
XX
DR   WPI; 2002-292254/33.
XX
PT   New derivatives of Bad peptide, useful for identifying compounds that
PT   bind to Bcl-2 proteins, potential agents for treating cancer and
PT   degenerative disease  -
XX
PS   Claim 4; Page 16; 31pp; English.
XX
CC   The present invention relates to new peptides that are derived from a
CC   wild-type human Bad peptide and are able to bind to a member of the
CC   Bcl-2 protein family. The peptides are useful, when labelled, in
CC   competitive/displacement assays for identifying substances that bind to
CC   members of the Bcl-2 family and may induce or suppress apoptosis so are
CC   potentially useful for treating cancer (inducers) or degenerative
CC   diseases or ischemic injury (suppressors). The peptides of the invention
CC   have high helix propensity, maintain the contacts of the wild-type Bad
CC   peptide and, compared with the Bad peptide, may have better physical
CC   properties, particularly solubility. The present sequence represents one
CC   of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC   from the BH3 domain of the human wild-type Bad peptide.
XX
SQ   Sequence   24 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 24;
  Best Local Similarity   100.0%;  Pred. No. 2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 11
AAU78627
ID   AAU78627 standard; Peptide; 24 AA.
```

XX
AC AAU78627;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human Bad peptide #27 which binds to a member of the Bcl-2 family.
XX
KW Human; Bad; Bcl-2; apoptosis; cancer; inducer; degenerative disease;
KW ischemic injury; suppressor; BH3 domain.
XX
OS Homo sapiens.
XX
PN WO200220568-A2.
XX
PD 14-MAR-2002.
XX
PF 04-SEP-2001; 2001WO-US27410.
XX
PR 06-SEP-2000; 2000US-0656399.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Fesik SW, Meadows RP, Joseph MK, Olejniczak ET, Petros AM;
PI Nettesheim DG, Swift KM, Matayoshi E, Zhang H;
XX
DR WPI; 2002-292254/33.
XX
PT New derivatives of Bad peptide, useful for identifying compounds that
PT bind to Bcl-2 proteins, potential agents for treating cancer and
PT degenerative disease -
XX
PS Claim 15; Page 18; 31pp; English.
XX
CC The present invention relates to new peptides that are derived from a
CC wild-type human Bad peptide and are able to bind to a member of the
CC Bcl-2 protein family. The peptides are useful, when labelled, in
CC competitive/displacement assays for identifying substances that bind to
CC members of the Bcl-2 family and may induce or suppress apoptosis so are
CC potentially useful for treating cancer (inducers) or degenerative
CC diseases or ischemic injury (suppressors). The peptides of the invention
CC have high helix propensity, maintain the contacts of the wild-type Bad
CC peptide and, compared with the Bad peptide, may have better physical
CC properties, particularly solubility. The present sequence represents one
CC of a collection of Bad peptides (AAU78601-AAU78631) that were derived
CC from the BH3 domain of the human wild-type Bad peptide.
XX
SQ Sequence 24 AA;

Query Match 100.0%; Score 83; DB 23; Length 24;
Best Local Similarity 100.0%; Pred. No. 2e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFVD 21

RESULT 12
ABP56161
ID ABP56161 standard; peptide; 25 AA.
XX
AC ABP56161;
XX
DT 28-MAR-2003 (first entry)
XX
DE PTPC-interacting TOX peptide #27.
XX
KW Mitochondrial membrane permeabilisation; mitochondrion; PTPC;
KW permeability transition pore complex; virucide; neuroprotective;
KW vasotropic; cytostatic; infection; cell death regulation; apoptosis;
KW mitochondrial permeability transition pore complex modulator; cancer;
KW apoptogenic; ischaemia; neurodegenerative disease; fulminant hepatitis.
XX
OS Synthetic.
XX
PN WO200261105-A2.
XX
PD 08-AUG-2002.
XX
PF 01-FEB-2002; 2002WO-EP01633.
XX
PR 02-FEB-2001; 2001US-265594P.
XX
PA (INSP) INST PASTEUR.
PA (CNRS) CENT NAT RECH SCI.
XX
PI Edelman L, Jacotot E, Briand J;
XX
DR WPI; 2002-619260/66.
XX
PT New chimeric bifunctional molecules that target specific cells and
PT regulate the apoptosis function of the permeability transition pore
PT complex of the mitochondria, useful for treating or preventing e.g.
PT cancer or ischemia -
XX
PS Claim 9; Page 11; 76pp; English.
XX
CC The present invention describes a chimeric bifunctional molecule (I)
CC comprising at least a first functional molecule covalently linked to a
CC second functional molecule, which is able to modulate the activity of
CC the permeability transition pore complex (PTPC) of the mitochondria.
CC (I) has the function of specifically targeting and entering a tissue
CC cell population. The second functional molecule has the function of
CC specifically targeting, and inducing or preventing the death of the
CC cells by apoptosis by regulating the opening or the closing of the PTPC
CC of the mitochondria or its fragment. (I) has virucide, neuroprotective,
CC vasotropic and cytostatic activities, and can be used as a mitochondrial
CC permeability transition pore complex (PTPC) modulator. (I) is useful for
CC treating or preventing a pathological infection or disease. (I) is also
CC useful for regulating cell death regulatory molecules, specifically the
CC apoptogenic function of the PTPC, for treating e.g. cancer, ischaemia,
CC neurodegenerative diseases, fulminant hepatitis or viral infections.
CC The present sequence represents a PTPC-interacting TOX peptide which is
CC given in the exemplification of the present invention.
XX
SQ Sequence 25 AA;

Query Match 100.0%; Score 83; DB 23; Length 25;
Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 QRYGRELRRMSDEFVD 16
||||||||||||||||
Db 6 QRYGRELRRMSDEFVD 21

RESULT 13
ABG78481
ID ABG78481 standard; Peptide; 25 AA.
XX
AC ABG78481;
XX
DT 15-NOV-2002 (first entry)
XX
DE Human Bcl2 fluorescein labelled peptide #1.
XX
KW Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200240530-A2.
XX
PD 23-MAY-2002.
XX
PF 15-NOV-2001; 2001WO-US45693.

```
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 15; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   25 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 14
ABG78484
ID   ABG78484 standard; Peptide; 25 AA.
XX
AC   ABG78484;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mutant Bcl2 competitive binding assay peptide #1.
XX
KW   Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200240530-A2.
XX
PD   23-MAY-2002.
XX
PF   15-NOV-2001; 2001WO-US45693.
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 17; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   25 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 25;
  Best Local Similarity   100.0%;  Pred. No. 2.1e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21


RESULT 15
ABG78485
ID   ABG78485 standard; Peptide; 25 AA.
XX
AC   ABG78485;
XX
DT   15-NOV-2002  (first entry)
XX
DE   Mutant Bcl2 competitive binding assay peptide #2.
XX
KW   Human; Bcl2; BclXl; programmed cell death; apoptosis; mutant; mutein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200240530-A2.
XX
PD   23-MAY-2002.
XX
PF   15-NOV-2001; 2001WO-US45693.
XX
PR   20-NOV-2000; 2000US-0716395.
XX
PA   (ABBO ) ABBOTT LAB.
XX
PI   Fesik SW,  Petros AM,  Yoon H,  Nettesheim DG;
XX
DR   WPI; 2002-490141/52.
XX
PT   New mutant Bcl-2 proteins derived from a wild type human Bcl-2 protein,
PT   useful in biological assays to identify substances that block the
PT   ability of Bcl-2 to inhibit programmed cell death or apoptosis  -
XX
PS   Example 2; Page 17; 36pp; English.
XX
CC   This invention relates to a novel mutant protein which is derived from
CC   a wild type human Bcl-2 protein. The mutant is created by replacing a
CC   sequence of amino acid residues comprising a flexible loop from the wild
CC   type Bcl-2 protein with an amino acid sequence comprising at least two
CC   acidic amino acids. The mutant Bcl-2 protein comprises a 166 residue
CC   shown in the specification. The invention also comprises an assay for
CC   identifying substances that bind to the Bcl-2 protein. The protein
CC   sequences of the invention are useful in biological assays to identify
CC   substances that block the ability of Bcl-2 to inhibit programmed cell
CC   death or apoptosis. The present sequence represents a human Bcl2
CC   peptide of the invention.
XX
SQ   Sequence   25 AA;

  Query Match             100.0%;  Score 83;  DB 23;  Length 25;
```

Best Local Similarity 100.0%; Pred. No. 2.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21
```

Search completed: September 15, 2003, 17:22:14
Job time : 23.6286 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:22:21 ; Search time 8.34286 Seconds
(without alignments)
81.144 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	166	1	US-08-665-617-2	Sequence 2, Appli
2	83	100.0	168	2	US-08-717-123-2	Sequence 2, Appli
3	83	100.0	168	3	US-08-985-335-1	Sequence 1, Appli
4	83	100.0	168	3	US-08-985-335-7	Sequence 7, Appli
5	83	100.0	168	3	US-09-410-372-1	Sequence 1, Appli
6	83	100.0	168	3	US-09-410-372-7	Sequence 7, Appli
7	83	100.0	168	4	US-09-375-257-2	Sequence 2, Appli
8	73	88.0	23	1	US-08-333-565-10	Sequence 10, Appl
9	73	88.0	23	2	US-08-661-479-10	Sequence 10, Appl
10	73	88.0	59	2	US-08-733-505A-55	Sequence 55, Appl
11	73	88.0	59	2	US-08-733-505A-56	Sequence 56, Appl
12	73	88.0	59	2	US-08-733-505A-57	Sequence 57, Appl
13	73	88.0	59	2	US-08-733-505A-58	Sequence 58, Appl
14	73	88.0	204	1	US-08-333-565-2	Sequence 2, Appli
15	73	88.0	204	2	US-08-661-479-2	Sequence 2, Appli
16	73	88.0	204	2	US-08-733-505A-1	Sequence 1, Appli
17	73	88.0	204	2	US-08-733-505A-12	Sequence 12, Appl
18	73	88.0	204	2	US-08-733-505A-13	Sequence 13, Appl
19	73	88.0	204	2	US-08-733-505A-14	Sequence 14, Appl
20	70	84.3	204	2	US-08-717-123-3	Sequence 3, Appli
21	70	84.3	204	4	US-09-375-257-3	Sequence 3, Appli
22	67	80.7	16	1	US-08-333-565-26	Sequence 26, Appl
23	67	80.7	16	2	US-08-661-479-26	Sequence 26, Appl
24	44	53.0	1125	4	US-09-252-991A-18729	Sequence 18729, A
25	42	50.6	11	2	US-08-733-505A-34	Sequence 34, Appl
26	42	50.6	11	2	US-08-706-741B-69	Sequence 69, Appl
27	42	50.6	11	2	US-08-924-695A-69	Sequence 69, Appl
28	42	50.6	876	1	US-08-785-429-2	Sequence 2, Appli
29	42	50.6	876	3	US-08-996-621-2	Sequence 2, Appli
30	42	50.6	888	4	US-09-134-001C-3032	Sequence 3032, Ap
31	41	49.4	284	4	US-09-328-352-6559	Sequence 6559, Ap
32	40	48.2	394	4	US-09-252-991A-31592	Sequence 31592, A
33	40	48.2	724	4	US-09-328-352-7710	Sequence 7710, Ap
34	39	47.0	333	4	US-09-252-991A-28443	Sequence 28443, A
35	38.5	46.4	1112	4	US-09-252-991A-27256	Sequence 27256, A
36	38	45.8	376	4	US-09-252-991A-26270	Sequence 26270, A
37	38	45.8	432	3	US-09-075-087-2	Sequence 2, Appli
38	38	45.8	432	3	US-09-472-971-1	Sequence 1, Appli
39	38	45.8	575	3	US-08-913-805A-2	Sequence 2, Appli
40	38	45.8	575	3	US-08-913-805A-10	Sequence 10, Appl
41	38	45.8	575	3	US-09-442-629-2	Sequence 2, Appli
42	38	45.8	575	3	US-09-442-629-10	Sequence 10, Appl
43	37	44.6	338	4	US-09-252-991A-30165	Sequence 30165, A
44	37	44.6	520	4	US-09-391-104-10	Sequence 10, Appl
45	36	43.4	80	4	US-09-328-352-8004	Sequence 8004, Ap

ALIGNMENTS

```
RESULT 1
US-08-665-617-2
; Sequence 2, Application US/08665617
; Patent No. 5663316
;  GENERAL INFORMATION:
;    APPLICANT:  Xudong, Yin
;    TITLE OF INVENTION:  Gene and Protein for Regulation of Cell Death
;    NUMBER OF SEQUENCES:  2
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Saliwanchik & Saliwanchik
;      STREET:  2421 N.W. 41st Street, Suite A-1
;      CITY:  Gainesville
;      STATE:  Florida
;      COUNTRY:  USA
;      ZIP:  32606
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/665,617
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Saliwanchik, David R.
;      REGISTRATION NUMBER:  31,794
;      REFERENCE/DOCKET NUMBER:  CL-8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (352) 375-8100
;      TELEFAX:  (352) 372-5800
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  166 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-665-617-2

  Query Match           100.0%;  Score 83;  DB 1;  Length 166;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        106 QRYGRELRRMSDEFVD 121


RESULT 2
```

```
US-08-717-123-2
; Sequence 2, Application US/08717123
; Patent No. 5965703
;  GENERAL INFORMATION:
;    APPLICANT:  Horne, William A.
;    APPLICANT:  Oltersdorf, Tilman
;    TITLE OF INVENTION:  Human BAD Polypeptides, Encoding Nucleic
;    TITLE OF INVENTION:  Acids and Methods of Use
;    NUMBER OF SEQUENCES:  15
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Campbell and Flores
;      STREET:  4370 La Jolla Village Drive, Suite 700
;      CITY:  San Diego
;      STATE:  California
;      COUNTRY:  United States
;      ZIP:  92122
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/717,123
;      FILING DATE:  20-SEP-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Campbell, Cathryn A.
;      REGISTRATION NUMBER:  31,815
;      REFERENCE/DOCKET NUMBER:  P-ID 1929
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (619) 535-9001
;      TELEFAX:  (619) 535-8949
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-717-123-2

  Query Match             100.0%;  Score 83;  DB 2;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 3
US-08-985-335-1
; Sequence 1, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-08-985-335-1

  Query Match             100.0%;  Score 83;  DB 3;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 4
US-08-985-335-7
; Sequence 7, Application US/08985335
; Patent No. 6080847
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/985,335
;      FILING DATE:  Filed Herewith
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
```

```
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-08-985-335-7

  Query Match            100.0%;  Score 83;  DB 3;  Length 168;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 5
US-09-410-372-1
; Sequence 1, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  SYNORAB01
;      CLONE:  358673
US-09-410-372-1

  Query Match            100.0%;  Score 83;  DB 3;  Length 168;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 6
US-09-410-372-7
; Sequence 7, Application US/09410372
; Patent No. 6281334
;  GENERAL INFORMATION:
;    APPLICANT:  Hillman, Jennifer L.
;    APPLICANT:  Yue, Henry
;    APPLICANT:  Lal, Preeti
;    APPLICANT:  Shah, Purvi
;    APPLICANT:  Corley, Neil C.
;    TITLE OF INVENTION:  PROTEINS ASSOCIATED WITH CELL
;    TITLE OF INVENTION:  PROLIFERATION
;    NUMBER OF SEQUENCES:  9
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Incyte Pharmaceuticals, Inc.
;      STREET:  3174 Porter Dr.
;      CITY:  Palo Alto
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94304
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette
;      COMPUTER:  IBM Compatible
;      OPERATING SYSTEM:  DOS
;      SOFTWARE:  FastSEQ for Windows Version 2.0
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/410,372
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/985,335
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Billings, Lucy J.
;      REGISTRATION NUMBER:  36,749
;      REFERENCE/DOCKET NUMBER:  PF-0421 US
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  650-855-0555
;      TELEFAX:  650-845-4166
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  168 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    IMMEDIATE SOURCE:
;      LIBRARY:  GenBank
;      CLONE:  1683637
US-09-410-372-7

  Query Match            100.0%;  Score 83;  DB 3;  Length 168;
  Best Local Similarity  100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 7
US-09-375-257-2
; Sequence 2, Application US/09375257
; Patent No. 6504022
; GENERAL INFORMATION:
;  APPLICANT:  Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION:  HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE:  480140.428D1
;  CURRENT APPLICATION NUMBER: US/09/375,257
```

```
;  CURRENT FILING DATE:  1999-08-16
;  NUMBER OF SEQ ID NOS:  15
;  SOFTWARE:  FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-375-257-2

  Query Match             100.0%;  Score 83;  DB 4;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 2.2e-07;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 8
US-08-333-565-10
; Sequence 10, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-333-565-10

  Query Match             88.0%;  Score 73;  DB 1;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db          8 QRYGRELRRMSDEF 21


RESULT 9
US-08-661-479-10
; Sequence 10, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  23 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-661-479-10

  Query Match             88.0%;  Score 73;  DB 2;  Length 23;
  Best Local Similarity   100.0%;  Pred. No. 1.5e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db          8 QRYGRELRRMSDEF 21


RESULT 10
US-08-733-505A-55
; Sequence 55, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
```

```
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  55:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-55

  Query Match             88.0%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 11
US-08-733-505A-56
; Sequence 56, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  56:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-56

  Query Match             88.0%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 12
US-08-733-505A-57
; Sequence 57, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  57:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-57

  Query Match             88.0%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 13
US-08-733-505A-58
; Sequence 58, Application US/08733505A
; Patent No. 5856445
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, STANLEY J.
;    TITLE OF INVENTION:  SERINE SUBSTITUTED MUTANTS OF
;    TITLE OF INVENTION:  BCL-XL/BCL-2 ASSOCIATED CELL DEATH REGULATOR
;    NUMBER OF SEQUENCES:  60
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  HOWELL & HAFERKAMP, L.C.
;      STREET:  7733 FORSYTH BLVD., SUITE 1400
;      CITY:  ST. LOUIS
;      STATE:  MISSOURI
;      COUNTRY:  USA
;      ZIP:  63105
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
```

```
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/733,505A
;      FILING DATE:
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  HOLLAND, DONALD R.
;      REGISTRATION NUMBER:  35,197
;      REFERENCE/DOCKET NUMBER:  965458
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (314) 727-5188
;      TELEFAX:  (314) 727-6092
;  INFORMATION FOR SEQ ID NO:  58:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  59 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  peptide
US-08-733-505A-58

  Query Match             88.0%;  Score 73;  DB 2;  Length 59;
  Best Local Similarity   100.0%;  Pred. No. 4.2e-06;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db         46 QRYGRELRRMSDEF 59


RESULT 14
US-08-333-565-2
; Sequence 2, Application US/08333565
; Patent No. 5622852
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/333,565
;      FILING DATE:  31-OCT-1994
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-333-565-2

  Query Match             88.0%;  Score 73;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        145 QRYGRELRRMSDEF 158


RESULT 15
US-08-661-479-2
; Sequence 2, Application US/08661479
; Patent No. 5834209
;  GENERAL INFORMATION:
;    APPLICANT:  KORSMEYER, Stanley J.
;    TITLE OF INVENTION:  Bcl-x/Bcl-2 ASSOCIATED CELL DEATH
;    TITLE OF INVENTION:  REGULATOR
;    NUMBER OF SEQUENCES:  59
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Townsend and Townsend Khourie and Crew
;      STREET:  379 Lytton Avenue
;      CITY:  Palo Alto
;      STATE:  California
;      COUNTRY:  US
;      ZIP:  94301
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/661,479
;      FILING DATE:  11-JUN-1995
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/333,565
;      FILING DATE:  31-OCT-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Smith, William M
;      REGISTRATION NUMBER:  30,223
;      REFERENCE/DOCKET NUMBER:  15726A-000700
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 326-2400
;      TELEFAX:  (415) 326-2422
;  INFORMATION FOR SEQ ID NO:  2:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  204 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
;    FEATURE:
;      NAME/KEY:  Protein
;      LOCATION:  1..204
;      OTHER INFORMATION:  /note= "Deduced amino acid sequence
;      OTHER INFORMATION:  of mouse BAD."
US-08-661-479-2

  Query Match             88.0%;  Score 73;  DB 2;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 1.6e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        145 QRYGRELRRMSDEF 158


Search completed: September 15, 2003, 17:45:06
Job time : 8.34286 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:25:56 ; Search time 12.6857 Seconds
(without alignments)
184.034 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 541936 seqs, 145912426 residues

Total number of hits satisfying chosen parameters: 541936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	25	15	US-10-059-261-258	Sequence 258, App
2	83	100.0	168	9	US-09-922-378-2	Sequence 2, Appli
3	83	100.0	168	9	US-09-894-657-1	Sequence 1, Appli
4	83	100.0	168	9	US-09-894-657-7	Sequence 7, Appli
5	83	100.0	168	14	US-10-066-179-2	Sequence 2, Appli
6	70	84.3	204	9	US-09-922-378-3	Sequence 3, Appli
7	70	84.3	204	14	US-10-066-179-3	Sequence 3, Appli
8	56	67.5	15	15	US-10-174-105A-147	Sequence 147, App
9	44	53.0	215	15	US-10-156-761-9145	Sequence 9145, Ap
10	42	50.6	647	15	US-10-156-761-9735	Sequence 9735, Ap
11	42	50.6	810	9	US-09-815-242-12636	Sequence 12636, A
12	42	50.6	876	9	US-09-815-242-13003	Sequence 13003, A
13	42	50.6	879	9	US-09-815-242-5312	Sequence 5312, Ap
14	41	49.4	312	15	US-10-156-761-7630	Sequence 7630, Ap
15	41	49.4	312	15	US-10-156-761-7867	Sequence 7867, Ap
16	40.5	48.8	1265	15	US-10-198-070-69	Sequence 69, Appl
17	40	48.2	113	9	US-09-815-242-13424	Sequence 13424, A
18	40	48.2	380	15	US-10-156-761-13594	Sequence 13594, A
19	39	47.0	454	15	US-10-166-087-46	Sequence 46, Appl
20	39	47.0	959	12	US-10-342-224-40	Sequence 40, Appl
21	38.5	46.4	588	15	US-10-156-761-10330	Sequence 10330, A
22	38	45.8	35	15	US-10-092-750-1	Sequence 1, Appli
23	38	45.8	50	10	US-09-971-980-64	Sequence 64, Appl
24	38	45.8	64	10	US-09-971-980-62	Sequence 62, Appl
25	38	45.8	138	15	US-10-092-750-241	Sequence 241, App
26	38	45.8	313	10	US-09-738-626-3724	Sequence 3724, Ap
27	38	45.8	320	10	US-09-738-626-4522	Sequence 4522, Ap
28	38	45.8	515	15	US-10-106-698-4658	Sequence 4658, Ap
29	38	45.8	575	9	US-09-839-136-2	Sequence 2, Appli
30	38	45.8	575	9	US-09-839-136-10	Sequence 10, Appl
31	37	44.6	391	15	US-10-163-198-19	Sequence 19, Appl
32	37	44.6	543	15	US-10-156-761-9820	Sequence 9820, Ap
33	37	44.6	556	12	US-09-949-029-36	Sequence 36, Appl
34	37	44.6	1270	15	US-10-223-070-8	Sequence 8, Appli
35	36	43.4	78	15	US-10-083-357-1237	Sequence 1237, Ap
36	36	43.4	80	15	US-10-106-698-5539	Sequence 5539, Ap
37	36	43.4	103	10	US-09-738-626-4453	Sequence 4453, Ap
38	36	43.4	270	11	US-09-934-455-162	Sequence 162, App
39	36	43.4	284	15	US-10-156-761-10335	Sequence 10335, A
40	36	43.4	292	10	US-09-738-626-4426	Sequence 4426, Ap
41	36	43.4	350	11	US-09-940-244-394	Sequence 394, App
42	36	43.4	393	10	US-09-950-510-22	Sequence 22, Appl
43	36	43.4	394	10	US-09-712-363-205	Sequence 205, App
44	36	43.4	404	15	US-10-156-761-14659	Sequence 14659, A
45	36	43.4	445	10	US-09-950-510-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-10-059-261-258
; Sequence 258, Application US/10059261
; Publication No. US20030077826A1
; GENERAL INFORMATION:
; APPLICANT: EDELMAN, LENA
; APPLICANT: JACOTOT, ETIENNE DANIEL FRANCOIS
; APPLICANT: BRIAND, JEAN-PAUL
; TITLE OF INVENTION: CHIMERIC MOLECULES CONTAINING A MODULE ABLE TO TARGET
; TITLE OF INVENTION: SPECIFIC CELLS AND A MODULE REGULATING THE APOPTOGENIC
; TITLE OF INVENTION: FUNCTION OF THE PERMEABILITY TRANSITION PORE COMPLEX
; TITLE OF INVENTION: (PTPC)
; FILE REFERENCE: 03495.0216
; CURRENT APPLICATION NUMBER: US/10/059,261
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 60/265,594
; PRIOR FILING DATE: 2001-02-02
; NUMBER OF SEQ ID NOS: 325
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 258
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: TOX peptide
US-10-059-261-258

Query Match 100.0%; Score 83; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 5.3e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db          6 QRYGRELRRMSDEFVD 21
```

RESULT 2
US-09-922-378-2

```
; Sequence 2, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-922-378-2

  Query Match            100.0%;  Score 83;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 3
US-09-894-657-1
; Sequence 1, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Yue, Henry
;                   Lal, Preeti
;                   Shah, Purvi
;                   Corley, Neil C.
;        TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                            PROLIFERATION
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/894,657
;             FILING DATE: 28-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/410,372
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0421 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 1:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 168 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: SYNORAB01
;             CLONE: 358673
;        SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-894-657-1

  Query Match            100.0%;  Score 83;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 4
US-09-894-657-7
; Sequence 7, Application US/09894657
; Patent No. US20020098569A1
;   GENERAL INFORMATION:
;        APPLICANT: Hillman, Jennifer L.
;                   Yue, Henry
;                   Lal, Preeti
;                   Shah, Purvi
;                   Corley, Neil C.
;        TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL
;                            PROLIFERATION
;        NUMBER OF SEQUENCES: 9
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: Incyte Pharmaceuticals, Inc.
;             STREET: 3174 Porter Dr.
;             CITY: Palo Alto
;             STATE: CA
;             COUNTRY: USA
;             ZIP: 94304
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Diskette
;             COMPUTER: IBM Compatible
;             OPERATING SYSTEM: DOS
;             SOFTWARE: FastSEQ for Windows Version 2.0
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/09/894,657
;             FILING DATE: 28-Jun-2001
;        PRIOR APPLICATION DATA:
;             APPLICATION NUMBER: 09/410,372
;             FILING DATE: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: Billings, Lucy J.
;             REGISTRATION NUMBER: 36,749
;             REFERENCE/DOCKET NUMBER: PF-0421 US
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 650-855-0555
;             TELEFAX: 650-845-4166
;   INFORMATION FOR SEQ ID NO: 7:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 168 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        IMMEDIATE SOURCE:
;             LIBRARY: GenBank
;             CLONE: 1683637
;        SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-894-657-7

  Query Match            100.0%;  Score 83;  DB 9;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 3.9e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 5
```

```
US-10-066-179-2
; Sequence 2, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 168
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-066-179-2

  Query Match            100.0%;  Score 83;  DB 14;  Length 168;
  Best Local Similarity  100.0%;  Pred. No. 3.9e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 6
US-09-922-378-3
; Sequence 3, Application US/09922378
; Patent No. US20020037869A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428D3
;  CURRENT APPLICATION NUMBER: US/09/922,378
;  CURRENT FILING DATE:  2001-08-03
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-922-378-3

  Query Match            84.3%;  Score 70;  DB 9;  Length 204;
  Best Local Similarity  92.9%;  Pred. No. 0.00064;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||:|||
Db        145 QRYGRELRRMTDEF 158


RESULT 7
US-10-066-179-3
; Sequence 3, Application US/10066179
; Publication No. US20020115631A1
; GENERAL INFORMATION:
;  APPLICANT: Horne, William A.
;  APPLICANT:  Oltersdorf, Tilman
;  TITLE OF INVENTION: HUMAN BAD POLYPEPTIDES, ENCODING NUCLEIC
;  TITLE OF INVENTION:  ACIDS AND METHODS OF USE
;  FILE REFERENCE: 480140.428C1
;  CURRENT APPLICATION NUMBER: US/10/066,179
;  CURRENT FILING DATE:  2002-02-01
;  NUMBER OF SEQ ID NOS: 15
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;   LENGTH: 204
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-10-066-179-3

  Query Match            84.3%;  Score 70;  DB 14;  Length 204;
  Best Local Similarity  92.9%;  Pred. No. 0.00064;
  Matches   13;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||:|||
Db        145 QRYGRELRRMTDEF 158


RESULT 8
US-10-174-105A-147
; Sequence 147, Application US/10174105A
; Publication No. US20030068652A1
; GENERAL INFORMATION:
;  APPLICANT: Cell Signaling Technology, Inc.
;  APPLICANT:  ZHANG, Hui
;  APPLICANT:  COMB, Michael J.
;  APPLICANT:  TAN, Yi
;  TITLE OF INVENTION: POSITIVE IDENTIFICATION OF PHOSPHO-PROTEINS USING MOTIF-SPECI
;  TITLE OF INVENTION:  CONTEXT-INDEPENDENT ANTIBODIES COUPLED WITH DATABASE SEARCHI
;  FILE REFERENCE: CST-138 CIP3
;  CURRENT APPLICATION NUMBER: US/10/174,105A
;  CURRENT FILING DATE:  2002-06-18
;  PRIOR APPLICATION NUMBER: US 09/148,712
;  PRIOR FILING DATE: 1998-09-04
;  PRIOR APPLICATION NUMBER: US 09/535,364
;  PRIOR FILING DATE: 2000-03-24
;  NUMBER OF SEQ ID NOS: 193
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 147
;   LENGTH: 15
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Synthetic Peptide
;   FEATURE:
;   NAME/KEY: MOD_RES
;   LOCATION: (8)..(8)
;   OTHER INFORMATION: PHOSPHORYLATION;  serine at position 8 is phosphorylated
US-10-174-105A-147

  Query Match            67.5%;  Score 56;  DB 15;  Length 15;
  Best Local Similarity  100.0%;  Pred. No. 0.0084;
  Matches   11;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEF 14
              ||||||||||
Db          1 GRELRRMSDEF 11


RESULT 9
US-10-156-761-9145
; Sequence 9145, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
```

```
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9145
;   LENGTH: 215
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9145

  Query Match             53.0%;  Score 44;  DB 15;  Length 215;
  Best Local Similarity   61.5%;  Pred. No. 12;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDE 13
              :|:| :|||| ||
Db        108 ERWGGDLRRMRDE 120


RESULT 10
US-10-156-761-9735
; Sequence 9735, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9735
;   LENGTH: 647
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-9735

  Query Match             50.6%;  Score 42;  DB 15;  Length 647;
  Best Local Similarity   57.1%;  Pred. No. 83;
  Matches    8;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              | | |||  || :|
Db        565 YARGLRRHGDELID 578


RESULT 11
US-09-815-242-12636
; Sequence 12636, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 12636
;   LENGTH: 810
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-12636

  Query Match             50.6%;  Score 42;  DB 9;  Length 810;
  Best Local Similarity   61.5%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFVD 16
              ||||  ::||:||
Db        185 GRELPILADEYVD 197


RESULT 12
US-09-815-242-13003
; Sequence 13003, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 13003
;   LENGTH: 876
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
US-09-815-242-13003

  Query Match             50.6%;  Score 42;  DB 9;  Length 876;
  Best Local Similarity   61.5%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;
```

```
Qy          4 GRELRRMSDEFVD 16
              ||||  ::||:||
Db        251 GRELPILADEYVD 263


RESULT 13
US-09-815-242-5312
; Sequence 5312, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
;  APPLICANT: Haselbeck, Robert
;  APPLICANT:  Ohlsen, Kari L.
;  APPLICANT:  Zyskind, Judith W.
;  APPLICANT:  Wall, Daniel
;  APPLICANT:  Trawick, John D.
;  APPLICANT:  Carr, Grant J.
;  APPLICANT:  Yamamoto, Robert T.
;  APPLICANT:  Xu, H. Howard
;  TITLE OF INVENTION: Identification of Essential Genes in
;  TITLE OF INVENTION:  Prokaryotes
;  FILE REFERENCE: ELITRA.011A
;  CURRENT APPLICATION NUMBER: US/09/815,242
;  CURRENT FILING DATE:  2001-03-21
;  PRIOR APPLICATION NUMBER: 60/191,078
;  PRIOR FILING DATE: 2000-03-21
;  PRIOR APPLICATION NUMBER: 60/206,848
;  PRIOR FILING DATE: 2000-05-23
;  PRIOR APPLICATION NUMBER: 60/207,727
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: 60/242,578
;  PRIOR FILING DATE: 2000-10-23
;  PRIOR APPLICATION NUMBER: 60/253,625
;  PRIOR FILING DATE: 2000-11-27
;  PRIOR APPLICATION NUMBER: 60/257,931
;  PRIOR FILING DATE: 2000-12-22
;  PRIOR APPLICATION NUMBER: 60/269,308
;  PRIOR FILING DATE: 2001-02-16
;  NUMBER OF SEQ ID NOS: 14110
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5312
;   LENGTH: 879
;   TYPE: PRT
;   ORGANISM: Staphylococcus aureus
;   FEATURE:
;   NAME/KEY: VARIANT
;   LOCATION: (1)...(879)
;   OTHER INFORMATION: Xaa = Any Amino Acid
US-09-815-242-5312

  Query Match             50.6%;  Score 42;  DB 9;  Length 879;
  Best Local Similarity   61.5%;  Pred. No. 1.1e+02;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFVD 16
              ||||  ::||:||
Db        255 GRELPILADEYVD 267


RESULT 14
US-10-156-761-7630
; Sequence 7630, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7630
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-7630

  Query Match             49.4%;  Score 41;  DB 15;  Length 312;
  Best Local Similarity   50.0%;  Pred. No. 57;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              |  | | ||: |:|:|
Db        125 QLSGEERRRLIDDFMD 140


RESULT 15
US-10-156-761-7867
; Sequence 7867, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
;  APPLICANT: OMURA, SATOSHI
;  APPLICANT:  IKEDA, HARUO
;  APPLICANT:  ISHIKAWA, JUN
;  APPLICANT:  HORIKAWA, HIROSHI
;  APPLICANT:  SHIBA, TADAYOSHI
;  APPLICANT:  SAKAKI, YOSHIYUKI
;  APPLICANT:  HATTORI, MASAHIRA
;  TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
;  FILE REFERENCE: 249-262
;  CURRENT APPLICATION NUMBER: US/10/156,761
;  CURRENT FILING DATE:  2002-05-29
;  PRIOR APPLICATION NUMBER: JP 2001-204089
;  PRIOR FILING DATE: 2001-05-30
;  PRIOR APPLICATION NUMBER: JP 2001-272697
;  PRIOR FILING DATE: 2001-08-02
;  NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 7867
;   LENGTH: 312
;   TYPE: PRT
;   ORGANISM: Streptomyces avermitilis
US-10-156-761-7867

  Query Match             49.4%;  Score 41;  DB 15;  Length 312;
  Best Local Similarity   50.0%;  Pred. No. 57;
  Matches    8;  Conservative    3;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              |  | | ||: |:|:|
Db        125 QLSGEERRRLIDDFMD 140


Search completed: September 15, 2003, 17:47:53
Job time : 13.6857 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:16:55 ; Search time 3.77143 Seconds
(without alignments)
199.507 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	83	100.0	168	1	BAD_HUMAN	Q92934 homo sapien
2	73	88.0	204	1	BAD_MOUSE	Q61337 mus musculu
3	73	88.0	205	1	BAD_RAT	O35147 rattus norv
4	42	50.6	196	1	BIM_MOUSE	O54918 mus musculu
5	42	50.6	196	1	BIM_RAT	O88498 rattus norv
6	42	50.6	457	1	HGD_RHILO	Q983j4 rhizobium l
7	42	50.6	653	1	HT2A_HUMAN	Q13049 homo sapien
8	41	49.4	370	1	AROG_CANAL	P79023 candida alb
9	41	49.4	861	1	GCR3_YEAST	P34160 saccharomyc
10	40	48.2	261	1	YGFG_ECOLI	P52045 escherichia
11	40	48.2	380	1	PHLC_TRYCR	O15886 trypanosoma
12	40	48.2	429	1	MTA1_ACEPA	O52702 acetobacter
13	40	48.2	631	1	RPSD_BORBU	P52323 borrelia bu
14	39	47.0	220	1	6PGL_THEMA	Q9x0n8 thermotoga
15	39	47.0	365	1	RA51_SCHPO	P36601 schizosacch
16	39	47.0	418	1	VSI2_REOVD	P03525 reovirus (t
17	39	47.0	418	1	VSI2_REOVL	P11314 reovirus (t
18	39	47.0	432	1	TIG_SALTY	Q8xfc4 salmonella
19	39	47.0	503	1	MTBR_BPRH1	P09915 bacteriopha
20	39	47.0	545	1	SYFB_METAC	Q8tpf7 methanosarc
21	39	47.0	545	1	SYFB_METMA	Q8pta5 methanosarc
22	39	47.0	880	1	SYV_BACST	P11931 bacillus st
23	39	47.0	1967	1	YG50_YEAST	P53327 saccharomyc
24	38.5	46.4	172	1	DEFL_RHIME	Q92qd0 rhizobium m
25	38.5	46.4	468	1	SELA_PSEAE	Q9hv01 pseudomonas
26	38	45.8	87	1	Y152_UREPA	Q9pqz2 ureaplasma
27	38	45.8	185	1	RRF_THEMA	Q9x1b9 thermotoga
28	38	45.8	198	1	BIM_HUMAN	O43521 homo sapien
29	38	45.8	251	1	KDKA_VIBCH	Q9kvb9 vibrio chol
30	38	45.8	384	1	ODP2_MYCGE	P47514 mycoplasma
31	38	45.8	402	1	ODP2_MYCPN	P75392 mycoplasma
32	38	45.8	432	1	TIG_ECOL6	Q8fka7 escherichia
33	38	45.8	432	1	TIG_ECOLI	P22257 escherichia
34	38	45.8	575	1	FUT8_BOVIN	Q9n0w2 b alpha-(1,
35	38	45.8	575	1	FUT8_HUMAN	Q9byc5 h alpha-(1,
36	38	45.8	575	1	FUT8_PIG	P79282 s alpha-(1,
37	38	45.8	631	1	EMF2_ARATH	Q8l6y4 arabidopsis
38	38	45.8	1521	1	EMB5_CAEEL	P34703 caenorhabdi
39	38	45.8	1535	1	LML1_CAEEL	Q18823 caenorhabdi
40	37	44.6	375	1	DP3B_MYCCA	P24117 mycoplasma
41	37	44.6	391	1	UBIF_ECOLI	P75728 escherichia
42	37	44.6	398	1	PSMR_ARCFU	O28303 archaeoglob
43	37	44.6	481	1	Y335_SYNY3	Q55587 synechocyst
44	37	44.6	777	1	BAR1_HUMAN	Q99728 homo sapien
45	37	44.6	787	1	RELA_MYCLE	Q49640 mycobacteri

ALIGNMENTS

RESULT 1
BAD_HUMAN
ID BAD_HUMAN STANDARD; PRT; 168 AA.
AC Q92934; O14803;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component 6) (Bcl-
DE XL/Bcl-2 associated death promoter) (BCL2-like 8 protein).
GN BAD OR BBC6 OR BCL2L8.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Yin D.X., Li Z., Huang B., Chen S., Zhou H.;
RT "A human protein that interacts with Bcl-2 and have homology to mouse
RT BAD.";
RL Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A., AND PHOSPHORYLATION BY RAF-1.
RX MEDLINE=97083574; PubMed=8929532;
RA Wang H.-G., Rapp U.R., Reed J.C.;
RT "Bcl-2 targets the protein kinase Raf-1 to mitochondria.";
RL Cell 87:629-638(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Takayama S., Reed J.C.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A., AND DIMERIZATION.
RC TISSUE=Bone marrow;
RX MEDLINE=98049554; PubMed=9388232;
RA Ottilie S., Diaz J.-L., Horne W., Chang J., Wang Y., Wilson G.,
RA Chang S., Weeks S., Fritz L.C., Oltersdorf T.;
RT "Dimerization properties of human BAD.";
RL J. Biol. Chem. 272:30866-30872(1997).
RN [5]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

```
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [6]
RP   STRUCTURE BY NMR OF 103-127.
RX   MEDLINE=21073561; PubMed=11206074;
RA   Petros A.M., Nettesheim D.G., Wang Y., Olejniczak E.T., Meadows R.P.,
RA   Mack J., Swift K., Matayoshi E.D., Zhang H., Thompson C.B.,
RA   Fesik S.W.;
RT   "Rationale for Bcl-xL/Bad peptide complex formation from structure,
RT   mutagenesis, and biophysical studies.";
RL   Protein Sci. 9:2528-2534(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-75/Ser-99 phosphorylated form binds 14-3-3 proteins (By
CC       similarity).
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A WIDE VARIETY OF TISSUES.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-75 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-99 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-118, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-99 is the major site of AKT/PKB phosphorylation, Ser-118 the
CC       major site of protein kinase A (CAPK) phosphorylation (by
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   -!- CAUTION: Ref.1 sequence differs from that shown due to frameshifts
CC       in position 64 and 91.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U66879; AAB36516.1; ALT_FRAME.
DR   EMBL; AF021792; AAB72092.1; -.
DR   EMBL; AF031523; AAB88124.1; -.
DR   EMBL; BC001901; AAH01901.1; -.
DR   PDB; 1G5J; 07-FEB-01.
DR   Genew; HGNC:936; BAD.
DR   MIM; 603167; -.
DR   GO; GO:0005737; C:cytoplasm; NAS.
DR   GO; GO:0005741; C:mitochondrial outer membrane; NAS.
DR   GO; GO:0005515; F:protein binding activity; NAS.
DR   GO; GO:0008632; P:apoptotic program; TAS.
DR   GO; GO:0006917; P:induction of apoptosis; NAS.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Polymorphism; 3D-structure.
FT   DOMAIN      110    124       BH3.
FT   MOD_RES      75     75       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES      99     99       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   MOD_RES     118    118       PHOSPHORYLATION (BY PKA AND PKB) (BY
FT                                SIMILARITY).
FT   VARIANT     107    107       A -> S (in dbSNP:3729933).
FT                                /FTId=VAR_015380.
FT   HELIX       106    121
SQ   SEQUENCE   168 AA;  18392 MW;  69FD8D27DDEE3241 CRC64;

  Query Match             100.0%;  Score 83;  DB 1;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 1.8e-06;
  Matches   16;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFVD 16
              ||||||||||||||||
Db        108 QRYGRELRRMSDEFVD 123


RESULT 2
BAD_MOUSE
ID   BAD_MOUSE      STANDARD;      PRT;   204 AA.
AC   Q61337;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD OR BBC6.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain, and Thymus;
RX   MEDLINE=95136361; PubMed=7834748;
RA   Yang E., Zha J., Jockel J., Boise L.H., Thompson C.B., Korsmeyer S.J.;
RT   "Bad, a heterodimeric partner for Bcl-XL and Bcl-2, displaces Bax and
RT   promotes cell death.";
RL   Cell 80:285-291(1995).
RN   [2]
RP   PHOSPHORYLATION, AND MUTAGENESIS OF SER-112 AND SER-136.
RX   MEDLINE=98022383; PubMed=9381178;
RA   Del Peso L., Gonzalez-Garcia M., Page C., Herrera R., Nunez G.;
RT   "Interleukin-3-induced phosphorylation of BAD through the protein
RT   kinase Akt.";
RL   Science 278:687-689(1997).
RN   [3]
RP   MUTAGENESIS OF SERINE RESIDUES.
RX   MEDLINE=20403302; PubMed=10949026;
RA   Datta S.R., Katsov A., Hu L., Petros A., Fesik S.W., Yaffe M.B.,
RA   Greenberg M.E.;
RT   "14-3-3 proteins and survival kinases cooperate to inactivate BAD by
RT   BH3 domain phosphorylation.";
RL   Mol. Cell 6:41-51(2000).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2.
CC       Appears to act as a link between growth factor receptor signaling
CC       and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10 (By similarity).
CC       The Ser-112/Ser-136 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-112 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-136 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-155, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-136 is the major site of AKT/PKB phosphorylation, Ser-155 the
CC       major site of protein kinase A (CAPK) phosphorylation.
```

```
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; L37296; AAA64465.1; -.
DR   PIR; A55671; A55671.
DR   HSSP; Q92934; 1G5J.
DR   MGD; MGI:1096330; Bad.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation.
FT   DOMAIN      147    161       BH3.
FT   MOD_RES     112    112       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     136    136       PHOSPHORYLATION (BY PKA AND PKB).
FT   MOD_RES     155    155       PHOSPHORYLATION (BY PKA AND PKB).
FT   MUTAGEN     112    112       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     136    136       S->A: NO PHOSPHORYLATION.
FT   MUTAGEN     155    155       S->A: NO PHOSPHORYLATION; INTERACTS WITH
FT                                BCL-X(L).
SQ   SEQUENCE   204 AA;  22080 MW;  6C2BA910205053F7 CRC64;

  Query Match             88.0%;  Score 73;  DB 1;  Length 204;
  Best Local Similarity   100.0%;  Pred. No. 9e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        145 QRYGRELRRMSDEF 158


RESULT 3
BAD_RAT
ID   BAD_RAT         STANDARD;      PRT;   205 AA.
AC   O35147; O70256; Q9JHX1;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Bcl2-antagonist of cell death (BAD) (Bcl-2 binding component
DE   6) (Bcl-xL/Bcl-2 associated death promoter).
GN   BAD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., AND MUTAGENESIS OF SER-113 AND SER-137.
RC   TISSUE=Ovary;
RX   MEDLINE=98034386; PubMed=9369453;
RA   Hsu S.Y., Kaipia A., Zhu L., Hsueh A.J.W.;
RT   "Interference of BAD (Bcl-xL/Bcl-2-associated death promoter)-induced
RT   apoptosis in mammalian cells by 14-3-3 isoforms and P11.";
RL   Mol. Endocrinol. 11:1858-1867(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=98194755; PubMed=9535132;
RA   D'Agata V., Magro G., Travali S., Musco S., Cavallaro S.;
RT   "Cloning and expression of the programmed cell death regulator BAD in
RT   the rat brain.";
RL   Neurosci. Lett. 243:137-140(1998).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).
RC   TISSUE=Brain;
RX   MEDLINE=21109372; PubMed=11161472;
RA   Hamner S., Arumae U., Yu L.-Y., Sun Y.-F., Saarma M., Lindholm D.;
RT   "Functional characterization of two splice variants of rat BAD and
RT   their interaction with Bcl-w in sympathetic neurons.";
RL   Mol. Cell. Neurosci. 17:97-106(2001).
CC   -!- FUNCTION: Promotes cell death. Successfully competes for the
CC       binding to Bcl-x(L), Bcl-2 and Bcl-w, thereby affecting the level
CC       of heterodimerization of these proteins with BAX. Can reverse the
CC       death repressor activity of Bcl-x(L), but not that of Bcl-2 (By
CC       similarity). Appears to act as a link between growth factor
CC       receptor signaling and the apoptotic pathways.
CC   -!- SUBUNIT: Forms heterodimers with the anti-apoptotic proteins, Bcl-
CC       x(L), Bcl-2 and Bcl-w. Also binds protein S100A10. The Ser-
CC       113/Ser-137 phosphorylated form binds 14-3-3 proteins.
CC   -!- SUBCELLULAR LOCATION: Outer mitochondrial membrane. Upon
CC       phosphorylation, locates to the cytoplasm (By similarity).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=2;
CC       Name=Alpha;
CC         IsoId=O35147-1; Sequence=Displayed;
CC       Name=Beta;
CC         IsoId=O35147-2; Sequence=VSP_000534;
CC   -!- TISSUE SPECIFICITY: Expressed in all tissues tested, including
CC       brain, liver, spleen and heart. In the brain, restricted to
CC       epithelial cells of the choroid plexus. Isoform alpha is the more
CC       abundant form.
CC   -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD AND
CC       BAX for their pro-apoptotic activity and for their interaction
CC       with anti-apoptotic members of the Bcl-2 family.
CC   -!- PTM: Phosphorylated on Ser-113 in response to survival stimuli.
CC       Subsequent phosphorylation on Ser-137 promotes heterodimerization
CC       with 14-3-3 proteins. This interaction then facilitates the
CC       phosphorylation at Ser-156, a site within the BH3 domain, leading
CC       to the release of Bcl-x(L) and the promotion of cell survival.
CC       Ser-137 is the major site of AKT/PKB phosphorylation, Ser-156 the
CC       major site of protein kinase A (CAPK) phosphorylation (By
CC       similarity).
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   -!- SIMILARITY: BELONGS TO THE BCL-2 FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AF003523; AAC53374.1; -.
DR   EMBL; AF031227; AAC15100.1; -.
DR   EMBL; AF279910; AAF91427.1; -.
DR   EMBL; AF279911; AAF91428.1; -.
DR   HSSP; Q92934; 1G5J.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Phosphorylation; Alternative splicing.
FT   DOMAIN      148    162       BH3.
FT   MOD_RES     113    113       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     137    137       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   MOD_RES     156    156       PHOSPHORYLATION (BY PKA AND PKB)
FT                                (BY SIMILARITY).
FT   VARSPLIC    166    205       LPRPKSAGTATQMRQSASWTRIIQSWWDRNLGKGGSTPSQ
FT                                -> EELTYSVEFLPVRAIAMEGWPLLWSFQSFPHTLPPTPP
FT                                EVAMFPLRYWTALRRLC (in isoform Beta).
FT                                /FTId=VSP_000534.
FT   MUTAGEN     113    113       S->A: NO EFFECT ON HETERODIMERIZATION
FT                                WITH 14-3-3 PROTEINS.
FT   MUTAGEN     137    137       S->A: NO HETERODIMERIZATION WITH 14-3-3
FT                                PROTEINS. NO EFFECT ON HETERODIMERIZATION
FT                                WITH BCL2 NOR WITH PROTEIN P11.
FT   CONFLICT     29     34       SDAGGR -> ERRGRK (IN REF. 1).
SQ   SEQUENCE   205 AA;  22228 MW;  7AFA71DAE9CF4A81 CRC64;

  Query Match             88.0%;  Score 73;  DB 1;  Length 205;
```

```
  Best Local Similarity   100.0%;  Pred. No. 9e-05;
  Matches   14;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        146 QRYGRELRRMSDEF 159


RESULT 4
BIM_MOUSE
ID   BIM_MOUSE      STANDARD;      PRT;   196 AA.
AC   O54918; O54919; O54920;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   BCL2-like protein 11 (BCL2 interacting mediator of cell death).
GN   BCL2L11 OR BIM.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBCELLULAR LOCATION, TISSUE
RP   SPECIFICITY, AND ALTERNATIVE SPLICING.
RX   MEDLINE=98094360; PubMed=9430630;
RA   O'Connor L., Strasser A., O'Reilly L.A., Hausmann G., Adams J.M.,
RA   Cory S., Huang D.C.S.;
RT   "Bim: a novel member of the Bcl-2 family that promotes apoptosis.";
RL   EMBO J. 17:384-395(1998).
CC   -!- FUNCTION: INDUCES APOPTOSIS. THE ISOFORMS VARY IN CYTOTOXICITY
CC       WITH ISOFORM BIMS BEING THE MOST POTENT AND ISOFORM BIMEL BEING
CC       THE LEAST POTENT.
CC   -!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
CC       PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
CC       NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
CC       BAX OR BAK (BY SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOPLASMIC MEMBRANES.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=BimEL;
CC         IsoId=O54918-1; Sequence=Displayed;
CC       Name=BimL;
CC         IsoId=O54918-2; Sequence=VSP_000536;
CC       Name=BimS;
CC         IsoId=O54918-3; Sequence=VSP_000537;
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN A NUMBER OF B-AND T-LYMPHOID CELL
CC       LINES.
CC   -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
CC       CYTOTOXICITY.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF032459; AAC40029.1; -.
DR   EMBL; AF032460; AAC40030.1; -.
DR   EMBL; AF032461; AAC40031.1; -.
DR   MGD; MGI:1197519; Bcl2l11.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Alternative splicing; Membrane.
FT   DOMAIN      146    160       BH3.
FT   VARSPLIC     42     97       Missing (in isoform BimL).
FT                                /FTId=VSP_000536.
FT   VARSPLIC     42    127       Missing (in isoform BimS).
FT                                /FTId=VSP_000537.
SQ   SEQUENCE   196 AA;  22066 MW;  531C176E5F1AC9AA CRC64;

  Query Match           50.6%;  Score 42;  DB 1;  Length 196;
  Best Local Similarity   61.5%;  Pred. No. 8.7;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEF 14
              |  :|||||: |||
Db        145 RIAQELRRIGDEF 157


RESULT 5
BIM_RAT
ID   BIM_RAT        STANDARD;      PRT;   196 AA.
AC   O88498; O88497; Q9WUI8;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   BCL2-like protein 11 (BCL2 interacting mediator of cell death)
DE   (Bcl-2 related ovarian death protein).
GN   BCL2L11 OR BIM OR BOD.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A., FUNCTION, SUBUNIT, AND TISSUE SPECIFICITY
RP   (ISOFORMS BOD-L; BOD-M AND BOD-S).
RC   TISSUE=Ovary;
RX   MEDLINE=98400436; PubMed=9731710;
RA   Hsu S.Y., Lin P., Hsueh A.J.W.;
RT   "BOD (Bcl-2-related ovarian death gene) is an ovarian BH3 domain-
RT   containing proapoptotic Bcl-2 protein capable of dimerization with
RT   diverse antiapoptotic Bcl-2 members.";
RL   Mol. Endocrinol. 12:1432-1440(1998).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM BIML).
RA   Chen D., Simon R.P., Chen J.;
RT   "Cloning of rat bimEL and bimL, and their differential expression in
RT   ischemia and normal rat brain.";
RL   Submitted (MAR-1999) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: INDUCES APOPTOSIS.
CC   -!- SUBUNIT: FORMS HETERODIMERS WITH A NUMBER OF ANTIAPOPTOTIC BCL-2
CC       PROTEINS INCLUDING MCL-1, BCL-2, BCL-XL, BFL-1, AND BHRF-1. DOES
CC       NOT HETERODIMERIZE WITH PROAPOPTOTIC PROTEINS SUCH AS BAD, BOK,
CC       BAX OR BAK.
CC   -!- SUBCELLULAR LOCATION: ASSOCIATED WITH INTRACYTOPLASMIC MEMBRANES
CC       (BY SIMILARITY).
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=3;
CC       Name=BOD-L;
CC         IsoId=O88498-1; Sequence=Displayed;
CC         Note=Isoform BOD-S is produced by alternative initation at
CC         Met-104 of isoform BOD-L;
CC       Name=BimL;
CC         IsoId=O88498-2; Sequence=VSP_000538;
CC       Name=BOD-M;
CC         IsoId=O88498-3; Sequence=VSP_000539;
CC       Event=Alternative initiation;
CC         Comment=2 isoforms, BOD-L (shown here) and BOD-S, are produced
CC         by alternative initiation at Met-1 and Met-104;
CC   -!- TISSUE SPECIFICITY: Widely expressed.
CC   -!- DOMAIN: THE BH3 DOMAIN IS REQUIRED FOR BCL-2 BINDING AND
CC       CYTOTOXICITY.
CC   -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF065433; AAC23595.1; -.
```

```
DR   EMBL; AF065431; AAC23593.1; -.
DR   EMBL; AF065432; AAC23594.1; -.
DR   EMBL; AF136927; AAD26594.1; -.
DR   InterPro; IPR000712; Bcl2_BH.
DR   PROSITE; PS01259; BH3; FALSE_NEG.
KW   Apoptosis; Alternative splicing; Membrane; Alternative initiation.
FT   CHAIN         1    196       BCL2-LIKE PROTEIN 11, ISOFORM BOD-L.
FT   CHAIN       104    196       BCL2-LIKE PROTEIN 11, ISOFORM BOD-S.
FT   INIT_MET    104    104       FOR ISOFORM BOD-S.
FT   DOMAIN      146    160       BH3.
FT   VARSPLIC     42     97       Missing (in isoform BimL).
FT                                /FTId=VSP_000538.
FT   VARSPLIC     42    127       Missing (in isoform BOD-M).
FT                                /FTId=VSP_000539.
FT   CONFLICT    136    136       E -> D (IN REF. 1; AAC23594).
SQ   SEQUENCE   196 AA;  22055 MW;  B4D2146F9C0B37A0 CRC64;

  Query Match             50.6%;  Score 42;  DB 1;  Length 196;
  Best Local Similarity   61.5%;  Pred. No. 8.7;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEF 14
              |  :||||: |||
Db        145 RIAQELRRIGDEF 157


RESULT 6
HGD_RHILO
ID   HGD_RHILO      STANDARD;      PRT;   457 AA.
AC   Q983J4;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Homogentisate 1,2-dioxygenase (EC 1.13.11.5) (Homogentisicase)
DE   (Homogentisate oxygenase) (Homogentisic acid oxidase).
GN   HMGA OR MLR8303.
OS   Rhizobium loti (Mesorhizobium loti).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Phyllobacteriaceae; Mesorhizobium.
OX   NCBI_TaxID=381;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MAFF303099;
RX   MEDLINE=21082930; PubMed=11214968;
RA   Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA   Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA   Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA   Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA   Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT   Mesorhizobium loti.";
RL   DNA Res. 7:331-338(2000).
CC   -!- CATALYTIC ACTIVITY: Homogentisate + O(2) = 4-maleylacetoacetate.
CC   -!- COFACTOR: Iron (By similarity).
CC   -!- PATHWAY: Catabolism of tyrosine; third step.
CC   -!- PATHWAY: Catabolism of phenylalanine; fourth step.
CC   -!- SIMILARITY: Belongs to the homogentisate dioxygenase family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AP003013; BAB53887.1; -.
DR   HAMAP; MF_00334; -; 1.
DR   InterPro; IPR005708; HmgA.
DR   Pfam; PF04209; HgmA; 1.
DR   TIGRFAMs; TIGR01015; hmgA; 1.
KW   Oxidoreductase; Dioxygenase; Phenylalanine catabolism;
KW   Tyrosine catabolism; Metal-binding; Iron; Complete proteome.
FT   METAL       351    351       IRON (BY SIMILARITY).
FT   METAL       357    357       IRON (BY SIMILARITY).
FT   METAL       387    387       IRON (BY SIMILARITY).
SQ   SEQUENCE   457 AA;  51046 MW;  6A20B69E9A2B2BD1 CRC64;

  Query Match             50.6%;  Score 42;  DB 1;  Length 457;
  Best Local Similarity   46.7%;  Pred. No. 21;
  Matches    7;  Conservative    2;  Mismatches    6;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFVD 16
              ||| ||    | ::|
Db        425 RYGAELETRQDNYID 439


RESULT 7
HT2A_HUMAN
ID   HT2A_HUMAN     STANDARD;      PRT;   653 AA.
AC   Q13049; Q9NQP8;
DT   01-NOV-1997 (Rel. 35, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Zinc-finger protein HT2A (72 kDa Tat-interacting protein) (Tripartite
DE   motif-containing protein 32).
GN   TRIM32 OR HT2A.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=95297135; PubMed=7778269;
RA   Fridell R.A., Harding L.S., Bogerd H.P., Cullen B.R.;
RT   "Identification of a novel human zinc finger protein that
RT   specifically interacts with the activation domain of lentiviral Tat
RT   proteins.";
RL   Virology 209:347-357(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Sehra H.;
RL   Submitted (MAY-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- FUNCTION: MAY PLAY A SIGNIFICANT ROLE IN MEDIATING THE BIOLOGICAL
CC       ACTIVITY OF THE HIV-1 TAT PROTEIN IN VIVO. BINDS SPECIFICALLY TO
CC       THE ACTIVATION DOMAIN OF HIV-1 TAT AND CAN ALSO INTERACT WITH THE
CC       HIV-2 AND EIAV TAT PROTEINS IN VIVO.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: SPLEEN, THYMUS, PROSTRATE, TESTIS, OVARY,
CC       INTESTINE AND COLON.
CC   -!- SIMILARITY: Contains 1 RING-type zinc finger.
CC   -!- SIMILARITY: Contains 1 B box-type zinc finger.
```

```
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U18543; AAA86474.1; -.
DR   EMBL; AL133284; CAB92723.1; -.
DR   EMBL; BC003154; AAH03154.1; -.
DR   HSSP; P29590; 1BOR.
DR   Genew; HGNC:16380; TRIM32.
DR   MIM; 602290; -.
DR   GO; GO:0005634; C:nucleus; TAS.
DR   GO; GO:0003713; F:transcription co-activator activity; TAS.
DR   InterPro; IPR001258; NHL.
DR   InterPro; IPR000315; Znf_Bbox.
DR   InterPro; IPR001841; Znf_ring.
DR   Pfam; PF01436; NHL; 5.
DR   Pfam; PF00643; zf-B_box; 1.
DR   Pfam; PF00097; zf-C3HC4; 1.
DR   SMART; SM00336; BBOX; 1.
DR   SMART; SM00184; RING; 1.
DR   PROSITE; PS50119; ZF_BBOX; 1.
DR   PROSITE; PS00518; ZF_RING_1; 1.
DR   PROSITE; PS50089; ZF_RING_2; 1.
KW   Zinc-finger; Nuclear protein.
FT   DOMAIN       2     6       POLY-ALA.
FT   ZN_FING     20    65       RING-TYPE.
FT   ZN_FING    103   133       B BOX-TYPE.
FT   CONFLICT    27    27       F -> I (IN REF. 1).
SQ   SEQUENCE   653 AA;  71988 MW;  D83B1595CA8378FD CRC64;

  Query Match             50.6%;  Score 42;  DB 1;  Length 653;
  Best Local Similarity   61.5%;  Pred. No. 30;
  Matches    8;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDE 13
              | || | ||: ||
Db        186 QEYGHEERRVQDE 198


RESULT 8
AROG_CANAL
ID   AROG_CANAL     STANDARD;      PRT;   370 AA.
AC   P79023;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Phospho-2-dehydro-3-deoxyheptonate aldolase, tyrosine-inhibited
DE   (EC 4.1.2.15) (Phospho-2-keto-3-deoxyheptonate aldolase) (DAHP
DE   synthetase) (3-deoxy-D-arabino-heptulosonate 7-phosphate synthase).
GN   ARO4.
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 11651 / B792;
RA   Sousa S., Pereira S.A., Livi G.P.;
RL   Submitted (APR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   PARTIAL SEQUENCE FROM N.A.
RX   MEDLINE=96207468; PubMed=8625423;
RA   Pereira S.A., Livi G.P.;
RT   "Aromatic amino-acid biosynthesis in Candida albicans: identification
RT   of the ARO4 gene encoding a second DAHP synthase.";
RL   Curr. Genet. 29:441-445(1996).
CC   -!- FUNCTION: STEREOSPECIFIC CONDENSATION OF PHOSPHOENOLPYRUVATE (PEP)
CC       AND D-ERYTHROSE-4-PHOSPHATE (E4P) GIVING RISE TO 3-DEOXY-D-
CC       ARABINO-HEPTULOSONATE-7-PHOSPHATE (DAHP).
CC   -!- CATALYTIC ACTIVITY: 2-dehydro-3-deoxy-D-arabino-heptonate 7-
CC       phosphate + phosphate = phosphoenolpyruvate + D-erythrose 4-
CC       phosphate + H(2)O.
CC   -!- ENZYME REGULATION: INHIBITED BY TYROSINE (BY SIMILARITY).
CC   -!- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC       first step.
CC   -!- SIMILARITY: BELONGS TO CLASS-I DAHP SYNTHETASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U53216; AAB48240.1; -.
DR   HSSP; P00886; 1QR7.
DR   InterPro; IPR006219; AroFGH.
DR   InterPro; IPR006218; DAHP1/KDSA.
DR   Pfam; PF00793; DAHP_synth_1; 1.
DR   ProDom; PD005060; AroFGH; 1.
DR   TIGRFAMs; TIGR00034; aroFGH; 1.
KW   Aromatic amino acid biosynthesis; Lyase; Multigene family.
SQ   SEQUENCE   370 AA;  40291 MW;  11E5E324C8D7B6DB CRC64;

  Query Match             49.4%;  Score 41;  DB 1;  Length 370;
  Best Local Similarity   42.9%;  Pred. No. 24;
  Matches    6;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              ||: |::::||  |
Db         77 YGKRLKKLADELKD 90


RESULT 9
GCR3_YEAST
ID   GCR3_YEAST     STANDARD;      PRT;   861 AA.
AC   P34160;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   GCR3 protein (STO1 protein) (SUT1 protein).
GN   GCR3 OR STO1 OR SUT1 OR YMR125W OR YM8564.07 OR YM9553.01.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92380925; PubMed=1512188;
RA   Uemura H., Jigami Y.;
RT   "GCR3 encodes an acidic protein that is required for expression of
RT   glycolytic genes in Saccharomyces cerevisiae.";
RL   J. Bacteriol. 174:5526-5532(1992).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Pandit S., Sternglanz R.;
RL   Submitted (DEC-1992) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RA   Guo Z., Russo P., Sherman F.;
RL   Submitted (XXX-1994) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288c / AB972;
RX   PubMed=9169872;
RA   Bowman S., Churcher C.M., Badcock K., Brown D., Chillingworth T.,
RA   Connor R., Dedman K., Devlin K., Gentles S., Hamlin N., Hunt S.,
RA   Jagels K., Lye G., Moule S., Odell C., Pearson D., Rajandream M.A.,
RA   Rice P., Skelton J., Walsh S., Whitehead S., Barrell B.G.;
RT   "The nucleotide sequence of Saccharomyces cerevisiae chromosome
```

```
RT   XIII.";
RL   Nature 387:90-93(1997).
CC   -!- FUNCTION: REQUIRED FOR EXPRESSION OF GLYCOLYTIC GENES. HAS
CC       CERTAIN CHARACTERISTICS OF A TRANSCRIPTIONAL ACTIVATOR.
CC   -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC   -!- SIMILARITY: SOME, TO HUMAN CBP80.
CC   -!- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN FROM POSITION 708
CC       ONWARD AND IS SHORTER (725 AA) DUE TO A FRAMESHIFT.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; D10224; BAA01076.1; ALT_SEQ.
DR   EMBL; L07650; -; NOT_ANNOTATED_CDS.
DR   EMBL; L27744; -; NOT_ANNOTATED_CDS.
DR   EMBL; Z49273; CAA89274.1; -.
DR   EMBL; Z48622; CAA88550.1; -.
DR   PIR; A44919; A44919.
DR   SGD; S0004732; STO1.
DR   GO; GO:0000243; C:commitment complex; IPI.
DR   GO; GO:0005846; C:snRNA cap binding complex; IDA.
DR   GO; GO:0003729; F:mRNA binding activity; IPI.
DR   GO; GO:0006371; P:mRNA splicing; IPI.
DR   InterPro; IPR003890; IF_eIF4G.
DR   Pfam; PF02854; MIF4G; 1.
DR   SMART; SM00543; MIF4G; 1.
KW   DNA-binding; Nuclear protein.
FT   DOMAIN       22     30       NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT   DOMAIN      774    801       ASP/GLU-RICH (ACIDIC).
FT   DOMAIN      802    825       ARG/LYS-RICH (BASIC).
FT   CONFLICT    164    164       D -> V (IN REF. 3).
FT   CONFLICT    633    633       R -> I (IN REF. 3).
FT   CONFLICT    704    704       A -> R (IN REF. 3).
SQ   SEQUENCE   861 AA;  100017 MW;  EDD04907BDC9207D CRC64;

  Query Match             49.4%;  Score 41;  DB 1;  Length 861;
  Best Local Similarity   40.0%;  Pred. No. 58;
  Matches    6;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFV 15
              :||  | | ::|:|:
Db        821 RRYSHEYRELADKFI 835


RESULT 10
YGFG_ECOLI
ID   YGFG_ECOLI     STANDARD;      PRT;   261 AA.
AC   P52045; P76643;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Hypothetical protein ygfG.
GN   YGFG OR B2919.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
CC   -!- SIMILARITY: BELONGS TO THE ENOYL-COA HYDRATASE/ISOMERASE FAMILY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; U28377; AAA69086.1; ALT_INIT.
DR   EMBL; AE000375; AAC75956.1; ALT_INIT.
DR   PDB; 1EF8; 24-MAY-00.
DR   PDB; 1EF9; 24-MAY-00.
DR   EcoGene; EG12972; ygfG.
DR   InterPro; IPR001753; EnCoA_hydrtse.
DR   Pfam; PF00378; ECH; 1.
DR   PROSITE; PS00166; ENOYL_COA_HYDRATASE; 1.
KW   Hypothetical protein; Lyase; Complete proteome; 3D-structure.
SQ   SEQUENCE   261 AA;  29172 MW;  B6A8A13EC2C2EBE0 CRC64;

  Query Match             48.2%;  Score 40;  DB 1;  Length 261;
  Best Local Similarity   50.0%;  Pred. No. 25;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              |||:|  :|  |:|
Db         21 YGRKLNALSKVFID 34


RESULT 11
PHLC_TRYCR
ID   PHLC_TRYCR     STANDARD;      PRT;   380 AA.
AC   O15886;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   Variant-surface-glycoprotein phospholipase C (EC 3.1.4.47) (VSG
DE   lipase) (Glycosylphosphatidylinositol-specific phospholipase C)
DE   (GPI-PLC).
OS   Trypanosoma cruzi.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX   NCBI_TaxID=5693;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Redpath M., Carnall N., Webb H., Courel M., Amorim A.,
RA   Cardosodealmeida M.L., Carrington M.;
RL   Submitted (JUL-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: BY HYDROLYSIS OF THE ATTACHED GLYCOLIPID, RELEASES
CC       SOLUBLE VARIANT SURFACE GLYCOPROTEIN CONTAINING PHOSPHOINOSITOL
CC       FROM THE CELL WALL OF T.BRUCEI AFTER CELL LYSIS. IT ALSO CLEAVES
CC       SIMILAR MEMBRANE ANCHORS ON SOME MAMMALIAN PROTEINS. VSG LIPASE
CC       MAY PLAY A ROLE IN PROCESSES SUCH AS PARASITE DIFFERENTIATION OR
CC       ANTIGENIC VARIATION (BY SIMILARITY).
CC   -!- CATALYTIC ACTIVITY: Variant-surface-glycoprotein 1,2-didecanoyl-
CC       sn-phosphatidylinositol + H(2)O = 1,2-didecanoylglycerol + soluble
CC       variant-surface-glycoprotein.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- SUBCELLULAR LOCATION: Membrane-associated.
CC   -!- SIMILARITY: DOMAIN X IS CONSERVED IN DIFFERENT FORMS OF PLC AND IS
CC       ESSENTIAL FOR CATALYTIC ACTIVITY.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; AJ000079; CAA03904.1; -.
DR   InterPro; IPR000909; PI_PLC_Xdom.
DR   InterPro; IPR003633; Varsurfglyc_PPLC.
DR   Pfam; PF00388; PI-PLC-X; 1.
```

```
DR   Pfam; PF03490; Varsurf_PPLC; 1.
DR   ProDom; PD041675; Varsurfglyc_PPLC; 1.
DR   SMART; SM00148; PLCXc; 1.
DR   PROSITE; PS50007; PIPLC_X_DOMAIN; 1.
KW   Hydrolase; Membrane.
FT   DOMAIN       31    205       DOMAIN X.
SQ   SEQUENCE   380 AA;  42736 MW;  273CD402B52068C5 CRC64;

  Query Match             48.2%;  Score 40;  DB 1;  Length 380;
  Best Local Similarity   50.0%;  Pred. No. 36;
  Matches    7;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFV 15
              :: ||| |:|| |:
Db        165 KFFRELDDRLSDRFI 178


RESULT 12
MTA1_ACEPA
ID   MTA1_ACEPA     STANDARD;      PRT;   429 AA.
AC   O52702;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   Modification methylase ApaLI (EC 2.1.1.73) (Cytosine-specific
DE   methyltransferase ApaLI) (M.ApaLI).
GN   APALIM.
OS   Acetobacter pasteurianus (Acetobacter turbidans).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC   Acetobacteraceae; Acetobacter.
OX   NCBI_TaxID=438;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 12875;
RX   MEDLINE=99077292; PubMed=9862476;
RA   Xu S.-Y., Xiao J.-P., Ettwiller L., Holden M., Aliotta J., Poh C.L.,
RA   Dalton M., Robinson D.P., Petronzio T.R., Moran L., Ganatra M.,
RA   Ware J., Slatko B., Benner J. II;
RT   "Cloning and expression of the ApaLI, NspI, NspHI, SacI, ScaI, and
RT   SapI restriction-modification systems in Escherichia coli.";
RL   Mol. Gen. Genet. 260:226-231(1998).
CC   -!- FUNCTION: THIS METHYLASE RECOGNIZES THE DOUBLE-STRANDED SEQUENCE
CC       GTGCAC, CAUSES SPECIFIC METHYLATION ON C-? ON BOTH STRANDS, AND
CC       PROTECTS THE DNA FROM CLEAVAGE BY THE APALI ENDONUCLEASE.
CC   -!- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + DNA cytosine = S-
CC       adenosyl-L-homocysteine + DNA 5-methylcytosine.
CC   -!- SIMILARITY: BELONGS TO THE C5-METHYLTRANSFERASE FAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF044847; AAC97180.1; -.
DR   REBASE; 3281; M.ApaLI.
DR   InterPro; IPR001525; C5_DNA_meth.
DR   Pfam; PF00145; DNA_methylase; 1.
DR   PRINTS; PR00105; C5METTRFRASE.
DR   TIGRFAMs; TIGR00675; dcm; 1.
DR   PROSITE; PS00094; C5_MTASE_1; FALSE_NEG.
KW   Transferase; Methyltransferase; Restriction system.
FT   ACT_SITE     81     81       BY SIMILARITY.
SQ   SEQUENCE   429 AA;  46547 MW;  E011C7D15B33F5F3 CRC64;

  Query Match             48.2%;  Score 40;  DB 1;  Length 429;
  Best Local Similarity   61.5%;  Pred. No. 41;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFVD 16
              ||:| |:  ||||
Db        129 GRDLARLVREFVD 141


RESULT 13
RPSD_BORBU
ID   RPSD_BORBU     STANDARD;      PRT;   631 AA.
AC   P52323;
DT   01-OCT-1996 (Rel. 34, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   RNA polymerase sigma factor rpoD (Sigma-70).
GN   RPOD OR BB0712.
OS   Borrelia burgdorferi (Lyme disease spirochete).
OC   Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE OF 89-631 FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RA   Pan M.;
RL   Thesis (1994), National Taiwan University, Taiwan.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 35210 / B31;
RX   MEDLINE=98065943; PubMed=9403685;
RA   Fraser C.M., Casjens S., Huang W.M., Sutton G.G., Clayton R.A.,
RA   Lathigra R., White O., Ketchum K.A., Dodson R., Hickey E.K., Gwinn M.,
RA   Dougherty B., Tomb J.-F., Fleischmann R.D., Richardson D.,
RA   Peterson J., Kerlavage A.R., Quackenbush J., Salzberg S., Hanson M.,
RA   van Vugt R., Palmer N., Adams M.D., Gocayne J.D., Weidman J.,
RA   Utterback T., Watthey L., McDonald L., Artiach P., Bowman C.,
RA   Garland S., Fujii C., Cotton M.D., Horst K., Roberts K., Hatch B.,
RA   Smith H.O., Venter J.C.;
RT   "Genomic sequence of a Lyme disease spirochaete, Borrelia
RT   burgdorferi.";
RL   Nature 390:580-586(1997).
RN   [3]
RP   SEQUENCE OF 165-614 FROM N.A.
RC   STRAIN=297;
RA   Pan M., Yeh J., Tsai C.;
RL   Submitted (SEP-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
CC       ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
CC       THEN IS RELEASED. THIS IS THE PRIMARY SIGMA-FACTOR OF THIS
CC       BACTERIA.
CC   -!- SIMILARITY: Belongs to the sigma-70 factor family.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; U17591; AAC44104.1; -.
DR   EMBL; AE001171; AAC67061.1; -.
DR   EMBL; U68006; AAC45100.1; -.
DR   PIR; G70188; G70188.
DR   HSSP; P00579; 1SIG.
DR   TIGR; BB0712; -.
DR   InterPro; IPR000943; Sigma_70.
DR   Pfam; PF03979; sigma70_r1_1; 1.
DR   Pfam; PF00140; sigma70_r1_2; 1.
DR   Pfam; PF04542; sigma70_r2; 1.
DR   Pfam; PF04539; sigma70_r3; 1.
DR   Pfam; PF04545; sigma70_r4; 1.
DR   PROSITE; PS00715; SIGMA70_1; 1.
DR   PROSITE; PS00716; SIGMA70_2; 1.
KW   Transcription regulation; Sigma factor; DNA-directed RNA polymerase;
KW   DNA-binding; Complete proteome.
FT   DOMAIN      419    432       POLYMERASE CORE BINDING (POTENTIAL).
FT   DNA_BIND    589    608       H-T-H MOTIF (BY SIMILARITY).
SQ   SEQUENCE   631 AA;  73642 MW;  BD565AB7D8F44796 CRC64;
```

```
  Query Match             48.2%;  Score 40;  DB 1;  Length 631;
  Best Local Similarity   50.0%;  Pred. No. 61;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              |  |||  ||:::|
Db        279 YQEELRIFSDDYID 292


RESULT 14
6PGL_THEMA
ID   6PGL_THEMA     STANDARD;      PRT;   220 AA.
AC   Q9X0N8;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   6-phosphogluconolactonase (EC 3.1.1.31) (6PGL).
GN   PGL OR DEVB OR TM1154.
OS   Thermotoga maritima.
OC   Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
OX   NCBI_TaxID=2336;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MSB8 / DSM 3109;
RX   MEDLINE=99287316; PubMed=10360571;
RA   Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA   Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA   McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA   Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA   Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA   Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT   "Evidence for lateral gene transfer between Archaea and Bacteria from
RT   genome sequence of Thermotoga maritima.";
RL   Nature 399:323-329(1999).
CC   -!- FUNCTION: HYDROLYSIS OF 6-PHOSPHOGLUCONOLACTONE TO 6-
CC       PHOSPHOGLUCONATE.
CC   -!- CATALYTIC ACTIVITY: 6-phospho-D-glucono-1,5-lactone + H(2)O = 6-
CC       phospho-D-gluconate.
CC   -!- PATHWAY: Pentose phosphate pathway; second step.
CC   -!- SIMILARITY: BELONGS TO THE GLUCOSAMINE/GALACTOSAMINE-6-PHOSPHATE
CC       ISOMERASE FAMILY. 6-PHOSPHOGLUCONOLACTONASE SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AE001772; AAD36230.1; -.
DR   PIR; F72289; F72289.
DR   TIGR; TM1154; -.
DR   InterPro; IPR006148; Gluc_gal_isom.
DR   InterPro; IPR005900; Phosphogluconlac.
DR   Pfam; PF01182; Glucosamine_iso; 1.
DR   TIGRFAMs; TIGR01198; pgl; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   220 AA;  25325 MW;  9B0FD07EE01E60C3 CRC64;

  Query Match             47.0%;  Score 39;  DB 1;  Length 220;
  Best Local Similarity   42.9%;  Pred. No. 30;
  Matches    6;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              ::| ||:|  :|:|
Db        113 EKYEREIRSATDQF 126


RESULT 15
RA51_SCHPO
ID   RA51_SCHPO     STANDARD;      PRT;   365 AA.
AC   P36601;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   DNA repair protein rhp51 (RAD51 homolog).
GN   RHP51 OR RAD51 OR SPAC644.14C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94051565; PubMed=8233794;
RA   Muris D.F.R., Vreeken K., Carr A.M., Broughton B.C., Lehmann A.R.,
RA   Lohman P.H.M., Pastink A.;
RT   "Cloning the RAD51 homologue of Schizosaccharomyces pombe.";
RL   Nucleic Acids Res. 21:4586-4591(1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93364417; PubMed=8358431;
RA   Shinohara A., Ogawa H., Matsuda Y., Ushio N., Ikeo K., Ogawa T.;
RT   "Cloning of human, mouse and fission yeast recombination genes
RT   homologous to RAD51 and recA.";
RL   Nat. Genet. 4:239-243(1993).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94252568; PubMed=8194753;
RA   Jang Y.K., Jin Y.H., Kim E.M., Hong S.H., Fabre F., Park S.D.;
RT   "Cloning and sequence analysis of rhp51+, a Schizosaccharomyces pombe
RT   homolog of the Saccharomyces cerevisiae RAD51 gene.";
RL   Gene 142:207-211(1994).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=972;
RX   MEDLINE=21848401; PubMed=11859360;
RA   Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA   Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA   Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA   Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA   Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA   Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA   James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA   Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA   Oliver K., O'Neil S., Pearson D., Quail M.A., Rabbinowitsch E.,
RA   Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA   Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA   Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA   Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA   Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA   Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,
RA   Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA   Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA   Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA   Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA   Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA   Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA   Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA   Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA   Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT   "The genome sequence of Schizosaccharomyces pombe.";
RL   Nature 415:871-880(2002).
CC   -!- FUNCTION: REQUIRED BOTH FOR RECOMBINATION AND FOR THE REPAIR OF
CC       DNA DAMAGE CAUSED BY X-RAYS.
CC   -!- SIMILARITY: BELONGS TO THE RECA FAMILY. RAD51 SUBFAMILY.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
```

```
CC   ---------------------------------------------------------------------
DR   EMBL; Z22691; CAA80399.1; -.
DR   EMBL; D13805; BAA02963.1; -.
DR   EMBL; Z24756; CAA80879.1; -.
DR   EMBL; Z24756; CAA80878.1; ALT_INIT.
DR   EMBL; AL355012; CAB90141.1; -.
DR   PIR; S42107; S42107.
DR   HSSP; Q06609; 1B22.
DR   GeneDB_SPombe; SPAC644.14c; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR000445; HhH.
DR   InterPro; IPR003583; HHH_1.
DR   InterPro; IPR001553; RecA.
DR   Pfam; PF00633; HHH; 1.
DR   ProDom; PD000229; RecA; 1.
DR   SMART; SM00382; AAA; 1.
DR   SMART; SM00278; HhH1; 1.
DR   PROSITE; PS50162; RECA_2; 1.
DR   PROSITE; PS50163; RECA_3; 1.
KW   DNA damage; DNA repair; ATP-binding; DNA recombination.
FT   NP_BIND     149    156       ATP (POTENTIAL).
FT   CONFLICT     15     15       T -> M (IN REF. 2 AND 4).
SQ   SEQUENCE   365 AA;  39823 MW;  9F26EB9FA4F3C2BA CRC64;

  Query Match             47.0%;  Score 39;  DB 1;  Length 365;
  Best Local Similarity   53.8%;  Pred. No. 50;
  Matches    7;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEF 14
              |: | |:|::|||
Db        269 RFMRTLQRLADEF 281


Search completed: September 15, 2003, 17:23:01
Job time : 4.77143 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:17:31 ; Search time 17.3714 Seconds
(without alignments)
237.680 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	59	71.1	146	13	Q9I9N2	Q9i9n2 brachydanio
2	46	55.4	458	5	Q8SV57	Q8sv57 encephalito
3	46	55.4	477	5	Q8SVF8	Q8svf8 encephalito
4	46	55.4	641	5	Q8SU90	Q8su90 encephalito
5	45	54.2	230	16	Q8XXS6	Q8xxs6 ralstonia s
6	45	54.2	564	16	Q9RUK9	Q9ruk9 deinococcus
7	44	53.0	247	2	Q9AJL5	Q9ajl5 streptomyce
8	44	53.0	726	16	Q8G3Z1	Q8g3z1 bifidobacte
9	44	53.0	1248	16	Q9HZQ3	Q9hzq3 pseudomonas
10	43	51.8	205	10	Q8L757	Q8l757 arabidopsis
11	43	51.8	804	16	Q8EW78	Q8ew78 mycoplasma
12	43	51.8	5635	5	Q9N9N1	Q9n9n1 leishmania
13	42	50.6	213	17	Q8TJ31	Q8tj31 methanosarc
14	42	50.6	260	16	Q8RCC4	Q8rcc4 thermoanaer
15	42	50.6	297	5	Q8MZ90	Q8mz90 drosophila
16	42	50.6	339	10	Q9FQ05	Q9fq05 atrichum an
17	42	50.6	548	2	P71029	P71029 burkholderi
18	42	50.6	592	5	Q9V7Y6	Q9v7y6 drosophila
19	42	50.6	592	5	Q960B3	Q960b3 drosophila
20	42	50.6	876	16	Q931Q1	Q931q1 staphylococ
21	42	50.6	876	16	Q99TJ8	Q99tj8 staphylococ
22	42	50.6	876	16	Q8CS83	Q8cs83 staphylococ
23	42	50.6	990	17	Q8TUT0	Q8tut0 methanopyru
24	41.5	50.0	191	16	Q8YG63	Q8yg63 brucella me
25	41.5	50.0	191	16	Q8G1Q8	Q8g1q8 brucella su
26	41	49.4	216	16	Q9KZC5	Q9kzc5 streptomyce
27	41	49.4	447	10	Q9M0B1	Q9m0b1 arabidopsis
28	41	49.4	497	10	Q8L517	Q8l517 arabidopsis
29	41	49.4	577	3	O59699	O59699 schizosacch
30	41	49.4	1303	12	Q9QBU0	Q9qbu0 potato mop-
31	41	49.4	2381	5	Q8IKD7	Q8ikd7 plasmodium
32	40.5	48.8	792	4	P78313	P78313 homo sapien
33	40.5	48.8	1224	4	P78311	P78311 homo sapien
34	40.5	48.8	1265	4	P78312	P78312 homo sapien
35	40	48.2	84	17	Q9HMU7	Q9hmu7 halobacteri
36	40	48.2	109	16	Q97SI9	Q97si9 streptococc
37	40	48.2	113	16	Q8DR57	Q8dr57 streptococc
38	40	48.2	164	17	Q8ZTI1	Q8zti1 pyrobaculum
39	40	48.2	219	17	Q9UZQ0	Q9uzq0 pyrococcus
40	40	48.2	259	17	Q8PUZ1	Q8puz1 methanosarc
41	40	48.2	275	16	Q8XD14	Q8xd14 escherichia
42	40	48.2	335	10	O64692	O64692 arabidopsis
43	40	48.2	360	16	Q9I0H8	Q9i0h8 pseudomonas
44	40	48.2	361	10	Q9FIC1	Q9fic1 arabidopsis
45	40	48.2	411	16	Q9A3L1	Q9a3l1 caulobacter

ALIGNMENTS

RESULT 1
Q9I9N2
ID Q9I9N2 PRELIMINARY; PRT; 146 AA.
AC Q9I9N2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bad.
GN BAD.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20373792; PubMed=10917738;
RA Inohara N., Nunez G.;
RT "Genes with homology to mammalian apoptosis regulators identified in
RT zebrafish.";
RL Cell Death Differ. 7:509-510(2000).
DR EMBL; AF231017; AAF66962.2; -.
DR HSSP; Q92934; 1G5J.
DR ZFIN; ZDB-GENE-000616-1; bad.
SQ SEQUENCE 146 AA; 16546 MW; 28A5650BB5107ECB CRC64;

Query Match 71.1%; Score 59; DB 13; Length 146;
Best Local Similarity 71.4%; Pred. No. 0.047;
Matches 10; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEF 14
              ::||::||||||||
Db         93 KKYGQQLRRMSDEF 106
```

RESULT 2
Q8SV57
ID Q8SV57 PRELIMINARY; PRT; 458 AA.
AC Q8SV57;

```
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein ECU06_1680.
GN   ECU06_1680.
OS   Encephalitozoon cuniculi.
OC   Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX   NCBI_TaxID=6035;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RA   Genoscope;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RX   MEDLINE=21576510; PubMed=11719806;
RA   Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA   Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P.,
RA   Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA   Weissenbach J., Vivares C.P.;
RT   "Genome sequence and gene compaction of the eukaryote parasite
RT   Encephalitozoon cuniculi.";
RL   Nature 414:450-453(2001).
DR   EMBL; AL590446; CAD25529.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   458 AA;  51914 MW;  24111BF78D4A534D CRC64;

  Query Match             55.4%;  Score 46;  DB 5;  Length 458;
  Best Local Similarity   46.7%;  Pred. No. 24;
  Matches    7;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFVD 16
              ||||:: || |: ::
Db         13 RYGRDVMRMLDDMIE 27


RESULT 3
Q8SVF8
ID   Q8SVF8      PRELIMINARY;      PRT;   477 AA.
AC   Q8SVF8;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein ECU06_0040.
GN   ECU06_0040.
OS   Encephalitozoon cuniculi.
OC   Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX   NCBI_TaxID=6035;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RA   Genoscope;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RX   MEDLINE=21576510; PubMed=11719806;
RA   Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA   Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P.,
RA   Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA   Weissenbach J., Vivares C.P.;
RT   "Genome sequence and gene compaction of the eukaryote parasite
RT   Encephalitozoon cuniculi.";
RL   Nature 414:450-453(2001).
DR   EMBL; AL590446; CAD25364.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   477 AA;  54039 MW;  B5DCF6299724CC96 CRC64;

  Query Match             55.4%;  Score 46;  DB 5;  Length 477;
  Best Local Similarity   46.7%;  Pred. No. 25;
  Matches    7;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFVD 16
              ||||:: || |: ::
Db         32 RYGRDVMRMLDDMIE 46


RESULT 4
Q8SU90
ID   Q8SU90      PRELIMINARY;      PRT;   641 AA.
AC   Q8SU90;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Hypothetical protein ECU11_0030.
GN   ECU11_0030.
OS   Encephalitozoon cuniculi.
OC   Eukaryota; Fungi; Microsporidia; Unikaryonidae; Encephalitozoon.
OX   NCBI_TaxID=6035;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RA   Genoscope;
RL   Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GB-M1;
RX   MEDLINE=21576510; PubMed=11719806;
RA   Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
RA   Prensier G., Barbe V., Peyretaillade E., Brottier P., Wincker P.,
RA   Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA   Weissenbach J., Vivares C.P.;
RT   "Genome sequence and gene compaction of the eukaryote parasite
RT   Encephalitozoon cuniculi.";
RL   Nature 414:450-453(2001).
DR   EMBL; AL590450; CAD25913.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   641 AA;  72408 MW;  A87E6EBE49732FC2 CRC64;

  Query Match             55.4%;  Score 46;  DB 5;  Length 641;
  Best Local Similarity   46.7%;  Pred. No. 34;
  Matches    7;  Conservative    5;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFVD 16
              ||||:: || |: ::
Db        203 RYGRDVMRMLDDMIE 217


RESULT 5
Q8XXS6
ID   Q8XXS6      PRELIMINARY;      PRT;   230 AA.
AC   Q8XXS6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Probable ATP-binding ABC transporter protein.
GN   RSC2037 OR RS03602.
OS   Ralstonia solanacearum (Pseudomonas solanacearum).
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Ralstoniaceae; Ralstonia.
OX   NCBI_TaxID=305;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GMI1000;
RX   MEDLINE=21681879; PubMed=11823852;
RA   Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA   Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
RA   Chandler M., Choisne N., Claudel-Renard C., Cunnac S., Demange N.,
RA   Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
RA   Siguier P., Thebault P., Whalen M., Wincker P., Levy M.,
RA   Weissenbach J., Boucher C.A.;
RT   "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL   Nature 415:497-502(2002).
DR   EMBL; AL646067; CAD15739.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
```

```
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
KW   Complete proteome.
SQ   SEQUENCE   230 AA;  25231 MW;  7C3FDA1E7A19A2A4 CRC64;

  Query Match           54.2%;  Score 45;  DB 16;  Length 230;
  Best Local Similarity   64.3%;  Pred. No. 17;
  Matches    9;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEF 14
              |  || |||: |||
Db        168 QEIGRTLRRLVDEF 181


RESULT 6
Q9RUK9
ID   Q9RUK9      PRELIMINARY;      PRT;   564 AA.
AC   Q9RUK9;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Glycosyl hydrolase, family 13.
GN   DR1375.
OS   Deinococcus radiodurans.
OC   Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC   Deinococcaceae; Deinococcus.
OX   NCBI_TaxID=1299;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=R1;
RX   MEDLINE=20036896; PubMed=10567266;
RA   White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA   Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA   Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA   Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,
RA   Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA   Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA   Fraser C.M.;
RT   "Genome sequence of the radioresistant bacterium Deinococcus
RT   radiodurans R1.";
RL   Science 286:1571-1577(1999).
DR   EMBL; AE001983; AAF10944.1; -.
DR   HSSP; P21332; 1UOK.
DR   TIGR; DR1375; -.
DR   InterPro; IPR006047; Alpha_amyl_cat.
DR   InterPro; IPR006589; Alp_amyl_cat_sub.
DR   Pfam; PF00128; alpha-amylase; 1.
DR   SMART; SM00642; Aamy; 1.
KW   Hydrolase; Complete proteome.
SQ   SEQUENCE   564 AA;  63667 MW;  B8F50B9B0DFC8D51 CRC64;

  Query Match           54.2%;  Score 45;  DB 16;  Length 564;
  Best Local Similarity   64.3%;  Pred. No. 44;
  Matches    9;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              | ||:||: ||| |
Db        283 YVREMRRVIDEFDD 296


RESULT 7
Q9AJL5
ID   Q9AJL5      PRELIMINARY;      PRT;   247 AA.
AC   Q9AJL5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   VarR.
GN   VARR.
OS   Streptomyces virginiae.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1961;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=21125565; PubMed=11222601;
RA   Namwat W., Lee C.K., Kinoshita H., Yamada Y., Nihira T.;
RT   "Identification of the varR gene as a transcriptional regulator of
RT   virginiamycin S resistance in Streptomyces virginiae.";
RL   J. Bacteriol. 183:2025-2031(2001).
CC   -!- SIMILARITY: BELONGS TO THE TETR/ACRR FAMILY OF TRANSCRIPTIONAL
CC       REGULATORS.
DR   EMBL; AB046994; BAB32408.1; -.
DR   HSSP; P09164; 2TCT.
DR   InterPro; IPR001647; HTH_TetR.
DR   InterPro; IPR004111; TetR_C.
DR   Pfam; PF00440; tetR; 1.
DR   Pfam; PF02909; tetR_C; 1.
DR   PRINTS; PR00455; HTHTETR.
KW   DNA-binding; Transcription; Transcription regulation.
SQ   SEQUENCE   247 AA;  27328 MW;  B6326F43374598A8 CRC64;

  Query Match           53.0%;  Score 44;  DB 2;  Length 247;
  Best Local Similarity   50.0%;  Pred. No. 27;
  Matches    7;  Conservative    5;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              ||:|: | :|||::
Db        186 YGKEVGRTADEFLE 199


RESULT 8
Q8G3Z1
ID   Q8G3Z1      PRELIMINARY;      PRT;   726 AA.
AC   Q8G3Z1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Narrowly conserved hypothetical protein.
GN   BL1609.
OS   Bifidobacterium longum.
OC   Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC   Bifidobacteriaceae; Bifidobacterium.
OX   NCBI_TaxID=216816;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=NCC 2705;
RX   MEDLINE=22294977; PubMed=12381787;
RA   Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B.,
RA   Pessi G., Zwahlen M.-C., Desiere F., Bork P., Delley M.,
RA   Pridmore R.D., Arigoni F.;
RT   "The genome sequence of Bifidobacterium longum reflects its adaptation
RT   to the human gastrointestinal tract.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR   EMBL; AE014794; AAN25398.1; -.
KW   Hypothetical protein; Complete proteome.
SQ   SEQUENCE   726 AA;  81866 MW;  6BEB86B61FCBC586 CRC64;

  Query Match           53.0%;  Score 44;  DB 16;  Length 726;
  Best Local Similarity   60.0%;  Pred. No. 85;
  Matches    9;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDEFVD 16
              | |   ||:||||:|
Db        241 RVGVVYRRLSDEFLD 255


RESULT 9
Q9HZQ3
ID   Q9HZQ3      PRELIMINARY;      PRT;  1248 AA.
AC   Q9HZQ3;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Cobalamin biosynthetic protein CobN.
GN   COBN OR PA2944.
OS   Pseudomonas aeruginosa.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=287;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ATCC 15692 / PAO1;
RX   MEDLINE=20437337; PubMed=10984043;
RA   Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrener P.,
RA   Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA   Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA   Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA   Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA   Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT   "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT   opportunistic pathogen.";
RL   Nature 406:959-964(2000).
DR   EMBL; AE004720; AAG06332.1; -.
DR   InterPro; IPR003672; CobN/Mg_chltase.
DR   InterPro; IPR000408; Reg_chr_condens.
DR   Pfam; PF02514; cobN-Mg_chel; 1.
DR   PROSITE; PS00626; RCC1_2; 1.
KW   Complete proteome.
SQ   SEQUENCE   1248 AA;  138499 MW;  C3D3DBFEE6736C7A CRC64;

  Query Match             53.0%;  Score 44;  DB 16;  Length 1248;
  Best Local Similarity   50.0%;  Pred. No. 1.5e+02;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 QRYG--RELRRMSDEFVD 16
              : ||  |:| |::||| |
Db        615 ESYGPLRDLERLADEFYD 632


RESULT 10
Q8L757
ID   Q8L757      PRELIMINARY;      PRT;   205 AA.
AC   Q8L757;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   AT4G02715.
OS   Arabidopsis thaliana (Mouse-ear cress).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC   eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX   NCBI_TaxID=3702;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA   Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA   Deng J.M., Hayashizaki Y., Hsuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA   Kawai J., Kim C.J., Narusaka M., Quach H.L., Sakurai T., Satou M.,
RA   Seki M., Shinn P., Tang C.C., Toroumi M., Wallender E.K., Wong C.,
RA   Wu H.C., Yamada K., Yu G., Yuan S., Shinozaki K., Ecker J.,
RA   Theologis A., Davis R.W.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Columbia;
RA   Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA   Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA   Hayashizaki Y., Shinozaki K.;
RT   "Arabidopsis thaliana full-length cDNA.";
RL   Submitted (NOV-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY136474; AAM97139.1; -.
DR   EMBL; AK118269; BAC42887.1; -.
KW   Hypothetical protein.
SQ   SEQUENCE   205 AA;  22583 MW;  1BD8D1358ECFFF81 CRC64;

  Query Match             51.8%;  Score 43;  DB 10;  Length 205;
  Best Local Similarity   61.5%;  Pred. No. 32;
  Matches    8;  Conservative    2;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFVD 16
              | ::|| | ||||
Db        100 GSQIRRCSSEFVD 112


RESULT 11
Q8EW78
ID   Q8EW78      PRELIMINARY;      PRT;   804 AA.
AC   Q8EW78;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Cation-transporting p-type ATPase.
GN   MYPE3250.
OS   Mycoplasma penetrans.
OC   Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX   NCBI_TaxID=28227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=HF-2;
RX   MEDLINE=22354719; PubMed=12466555;
RA   Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
RA   Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
RT   "The complete genomic sequence of Mycoplasma penetrans, an
RT   intracellular bacterial pathogen in humans.";
RL   Nucleic Acids Res. 30:5293-5300(2002).
DR   EMBL; AP004171; BAC44118.1; -.
KW   Complete proteome.
SQ   SEQUENCE   804 AA;  88008 MW;  69C71AA628FEF7A3 CRC64;

  Query Match             51.8%;  Score 43;  DB 16;  Length 804;
  Best Local Similarity   66.7%;  Pred. No. 1.4e+02;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFV 15
              | ||| ||||::
Db        428 GSELREMSDEYL 439


RESULT 12
Q9N9N1
ID   Q9N9N1      PRELIMINARY;      PRT;  5635 AA.
AC   Q9N9N1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Dynein heavy chain, cytosolic.
GN   L3302.02.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RA   Hilbert H., Wedler H., Wedler E., Duesterhoeft A., Ivens A.C.,
RA   Quail M., Rajandream M.A., Barrell B.G.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Friedlin;
RX   MEDLINE=98146435; PubMed=9477341;
RA   Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA   Smith D.F.;
RT   "A physical map of the Leishmania major Friedlin genome.";
RL   Genome Res. 8:135-145(1998).
DR   EMBL; AL359781; CAB95305.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR004273; Dynein_heavy.
DR   InterPro; IPR000169; SHprot_acsite.
```

```
DR   Pfam; PF03028; Dynein_heavy; 1.
DR   SMART; SM00382; AAA; 4.
DR   PROSITE; PS00639; THIOL_PROTEASE_HIS; 1.
KW   ATP-binding.
SQ   SEQUENCE   5635 AA;  620050 MW;  64A9EB81A9B14641 CRC64;

  Query Match             51.8%;  Score 43;  DB 5;  Length 5635;
  Best Local Similarity   72.7%;  Pred. No. 1.1e+03;
  Matches    8;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMS 11
              ||:||:||| |
Db       1535 QRFGRDLRRWS 1545


RESULT 13
Q8TJ31
ID   Q8TJ31      PRELIMINARY;      PRT;   213 AA.
AC   Q8TJ31;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ABC transporter, ATP-binding protein.
GN   MA3957.
OS   Methanosarcina acetivorans.
OC   Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
OC   Methanosarcinaceae; Methanosarcina.
OX   NCBI_TaxID=2214;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C2A / ATCC 35395 / DSM 2834;
RX   MEDLINE=21929760; PubMed=11932238;
RA   Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
RA   FitzHugh W., Calvo S., Engels R., Smirnov S., Atnoor D., Brown A.,
RA   Allen N., Naylor J., Stange-Thomann N., DeArellano K., Johnson R.,
RA   Linton L., McEwan P., McKernan K., Talamas J., Tirrell A., Ye W.,
RA   Zimmer A., Barber R.D., Cann I., Graham D.E., Grahame D.A., Guss A.M.,
RA   Hedderich R., Ingram-Smith C., Kuettner H.C., Krzycki J.A.,
RA   Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
RA   Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
RA   Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
RA   Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
RA   Metcalf W.W., Birren B.;
RT   "The genome of Methanosarcina acetivorans reveals extensive metabolic
RT   and physiological diversity.";
RL   Genome Res. 12:532-542(2002).
DR   EMBL; AE011107; AAM07308.1; -.
DR   InterPro; IPR003593; AAA_ATPase.
DR   InterPro; IPR003439; ABC_transporter.
DR   Pfam; PF00005; ABC_tran; 1.
DR   ProDom; PD000006; ABC_transporter; 1.
DR   SMART; SM00382; AAA; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
KW   ATP-binding; Complete proteome.
SQ   SEQUENCE   213 AA;  23676 MW;  B548F93D94D109C9 CRC64;

  Query Match             50.6%;  Score 42;  DB 17;  Length 213;
  Best Local Similarity   80.0%;  Pred. No. 49;
  Matches    8;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDE 13
              ||:| |||||
Db         59 GRDLHRMSDE 68


RESULT 14
Q8RCC4
ID   Q8RCC4      PRELIMINARY;      PRT;   260 AA.
AC   Q8RCC4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   ABC-type amino acid transport system, periplasmic component.
GN   ARTI OR TTE0512.
OS   Thermoanaerobacter tengcongensis.
OC   Bacteria; Firmicutes; Clostridia; Thermoanaerobacteriales;
OC   Thermoanaerobacteriaceae; Thermoanaerobacter.
OX   NCBI_TaxID=119072;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MB4 / JCM 11007;
RX   MEDLINE=21992816; PubMed=11997336;
RA   Bao Q., Tian Y., Li W., Xu Z., Xuan Z., Hu S., Dong W., Yang J.,
RA   Chen Y., Xue Y., Xu Y., Lai X., Huang L., Dong X., Ma Y., Ling L.,
RA   Tan H., Chen R., Wang J., Yu J., Yang H.;
RT   "A complete sequence of T. tengcongensis genome.";
RL   Genome Res. 12:689-700(2002).
DR   EMBL; AE013022; AAM23788.1; -.
DR   InterPro; IPR003439; ABC_transporter.
DR   InterPro; IPR001311; SBP/glu_receptor.
DR   InterPro; IPR001638; SBP_bac_3.
DR   Pfam; PF00497; SBP_bac_3; 1.
DR   SMART; SM00062; PBPb; 1.
DR   PROSITE; PS00211; ABC_TRANSPORTER; 1.
DR   PROSITE; PS01039; SBP_BACTERIAL_3; 1.
KW   Complete proteome.
SQ   SEQUENCE   260 AA;  28849 MW;  F3D87ED16B18DCBC CRC64;

  Query Match             50.6%;  Score 42;  DB 16;  Length 260;
  Best Local Similarity   58.3%;  Pred. No. 61;
  Matches    7;  Conservative    4;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          5 RELRRMSDEFVD 16
              ::| |:||||:|
Db        170 KQLNRVSDEFMD 181


RESULT 15
Q8MZ90
ID   Q8MZ90      PRELIMINARY;      PRT;   297 AA.
AC   Q8MZ90;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   AT26020p.
GN   CG30456 OR CG15612 OR CG15613.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA   Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA   George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA   Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA   Patel S., Phouanenavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA   Celniker S.;
RL   Submitted (MAY-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY113300; AAM29305.1; -.
DR   FlyBase; FBgn0050456; CG30456.
DR   InterPro; IPR000219; RhoGEF.
DR   Pfam; PF00621; RhoGEF; 1.
DR   PROSITE; PS50010; DH_2; 1.
SQ   SEQUENCE   297 AA;  35804 MW;  58FDBF3DF688D99A CRC64;

  Query Match             50.6%;  Score 42;  DB 5;  Length 297;
  Best Local Similarity   53.3%;  Pred. No. 70;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFV 15
              ||  |:|:|: |||:
Db          3 QRNRRDLKRLFDEFL 17
```

Search completed: September 15, 2003, 17:25:49
Job time : 18.3714 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: September 15, 2003, 17:18:16 ; Search time 7.2 Seconds
(without alignments)
213.708 Million cell updates/sec

Title: US-09-544-664-29
Perfect score: 83
Sequence: 1 QRYGRELRRMSDEFVD 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	73	88.0	204	2	A55671	bad protein - mous
2	45	54.2	564	2	H75403	glycosyl hydrolase
3	44	53.0	1248	2	G83278	cobalamin biosynth
4	42	50.6	876	2	E89949	valine-tRNA ligase
5	41.5	50.0	191	2	AD3414	hypothetical cytos
6	41	49.4	447	2	F85356	hypothetical prote
7	41	49.4	577	2	T40297	membrane transport
8	41	49.4	858	2	A44919	GCR3 protein - yea
9	40	48.2	84	2	F84388	hypothetical prote
10	40	48.2	109	2	B95043	conserved hypothet
11	40	48.2	113	2	D97913	conserved hypothet
12	40	48.2	219	2	A75088	hypothetical prote
13	40	48.2	275	2	E91102	probable enzyme [i
14	40	48.2	275	2	A85948	probable enzyme yg
15	40	48.2	275	2	F65076	hypothetical prote
16	40	48.2	335	2	T52577	gibberellin 2beta-
17	40	48.2	360	2	B83311	hypothetical prote
18	40	48.2	411	2	F87644	transcription regu
19	40	48.2	429	2	JC4986	site-specific DNA-
20	40	48.2	445	2	G97123	probable Fe-S oxid
21	40	48.2	631	2	G70188	transcription init
22	40	48.2	5138	2	B96695	hypothetical prote
23	39.5	47.6	198	2	E87441	conserved hypothet
24	39.5	47.6	414	2	B84275	hypothetical prote
25	39	47.0	73	2	AC3365	hypothetical prote
26	39	47.0	172	2	B71339	probable cationic
27	39	47.0	207	2	B95348	hypothetical prote
28	39	47.0	220	2	F72289	oxidoreductase, so
29	39	47.0	275	2	C69808	transporter homolo
30	39	47.0	331	2	E90121	DNA repair protein
31	39	47.0	360	2	D86200	protein F12K11.20
32	39	47.0	365	2	S42107	RAD51 protein homo
33	39	47.0	380	2	T32163	hypothetical prote
34	39	47.0	383	2	T31738	hypothetical prote
35	39	47.0	418	1	FOXRL2	sigma 2 protein -
36	39	47.0	418	1	FOXR3D	sigma 2 protein -
37	39	47.0	432	2	AB0558	trigger factor [im
38	39	47.0	503	1	CTBPRH	site-specific DNA-
39	39	47.0	536	2	F90299	acylaminoacyl-pept
40	39	47.0	689	2	T29772	hypothetical prote
41	39	47.0	880	1	SYBSVS	valine-tRNA ligase
42	39	47.0	959	1	B71405	probable kinesin -
43	39	47.0	965	2	AE2452	two-component hybr
44	39	47.0	1805	2	T02712	similar to late em
45	39	47.0	1967	2	S64604	hypothetical prote

ALIGNMENTS

RESULT 1
A55671
bad protein - mouse
C;Species: Mus musculus (house mouse)
C;Date: 03-Mar-1995 #sequence_revision 03-Mar-1995 #text_change 05-Nov-1999
C;Accession: A55671
R;Yang, E.; Zha, J.; Jockel, J.; Boise, L.H.; Thompson, C.B.; Korsmeyer, S.J.
Cell 80, 285-291, 1995
A;Title: Bad, a heterodimeric partner for Bcl-x-L and Bcl-2, displaces Bax and promot
A;Reference number: A55671; MUID:95136361; PMID:7834748
A;Accession: A55671
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-204 <YAN>
A;Cross-references: GB:L37296; NID:g639778; PIDN:AAA64465.1; PID:g639779
C;Keywords: heterodimer

Query Match 88.0%; Score 73; DB 2; Length 204;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 QRYGRELRRMSDEF 14
              ||||||||||||||
Db        145 QRYGRELRRMSDEF 158
```

RESULT 2
H75403
glycosyl hydrolase, family 13 - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C;Accession: H75403
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.
, M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: H75403
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-564 <WHI>
A;Cross-references: GB:AE001983; GB:AE000513; NID:g6459123; PIDN:AAF10944.1; PID:g64
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR1375
A;Map position: 1
C;Superfamily: alpha-glucosidase; alpha-amylase core homology

Query Match 54.2%; Score 45; DB 2; Length 564;
Best Local Similarity 64.3%; Pred. No. 14;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

```
Qy          3 YGRELRRMSDEFVD 16
              | ||:||: ||| |
Db        283 YVREMRRVIDEFDD 296


RESULT 3
G83278
cobalamin biosynthetic protein CobN PA2944 [imported] - Pseudomonas aeruginosa (strain P
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 02-Mar-2001
C;Accession: G83278
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: G83278
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1248 <STO>
A;Cross-references: GB:AE004720; GB:AE004091; NID:g9949032; PIDN:AAG06332.1; GSPDB:GN001
A;Experimental source: strain PA01
C;Genetics:
A;Gene: cobN; PA2944
C;Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

  Query Match             53.0%;  Score 44;  DB 2;  Length 1248;
  Best Local Similarity   50.0%;  Pred. No. 47;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          1 QRYG--RELRRMSDEFVD 16
              : ||  |:| |::||| |
Db        615 ESYGPLRDLERLADEFYD 632


RESULT 4
E89949
valine-tRNA ligase [imported] - Staphylococcus aureus (strain N315)
C;Species: Staphylococcus aureus
C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C;Accession: E89949
R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
 C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A;Title: Whole genome sequencing of meticillin-resistant Stapylococcus aureus.
A;Reference number: A89758; MUID:21311952; PMID:11418146
A;Accession: E89949
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-876 <KUR>
A;Cross-references: GB:BA000018; PID:g13701460; PIDN:BAB42754.1; GSPDB:GN00149
A;Experimental source: strain N315
C;Genetics:
A;Gene: valS
C;Superfamily: valine-tRNA ligase

  Query Match             50.6%;  Score 42;  DB 2;  Length 876;
  Best Local Similarity   61.5%;  Pred. No. 70;
  Matches    8;  Conservative    3;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          4 GRELRRMSDEFVD 16
              ||||  ::||:||
Db        251 GRELPILADEYVD 263


RESULT 5
AD3414
hypothetical cytosolic protein BMEI1298 [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AD3414
R;DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
.; Mazur, M.; Goltsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melit
A;Reference number: AD3252; PMID:11756688
A;Accession: AD3414
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-191 <KUR>
A;Cross-references: GB:AE008917; PIDN:AAL52479.1; PID:g17983287; GSPDB:GN00190
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI1298
A;Map position: I

  Query Match             50.0%;  Score 41.5;  DB 2;  Length 191;
  Best Local Similarity   33.3%;  Pred. No. 18;
  Matches   11;  Conservative    2;  Mismatches    3;  Indels   17;  Gaps    1;

Qy          1 QRYGR----------------ELRRMSDEFVD 16
              || ||                |||| :| |:|
Db        131 QRKGRKVSVVSTLTTQPAMISDELRRQADHFID 163


RESULT 6
F85356
hypothetical protein AT4g30490 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C;Accession: F85356
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A;Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A;Reference number: A85001; MUID:20083488; PMID:10617198
A;Accession: F85356
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-447 <STO>
A;Cross-references: GB:NC_001268; NID:g7269950; PIDN:CAB79767.1; GSPDB:GN00140
C;Genetics:
A;Gene: AT4g30490
A;Map position: 4

  Query Match             49.4%;  Score 41;  DB 2;  Length 447;
  Best Local Similarity   66.7%;  Pred. No. 52;
  Matches    8;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          5 RELRRMSDEFVD 16
              |||:|: || ||
Db         84 RELQRLYDELVD 95


RESULT 7
T40297
membrane transporter - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 18-Feb-2000
C;Accession: T40297
R;Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hilbert, H.; Moestl, D.; Duest
submitted to the EMBL Data Library, May 1998
A;Reference number: Z21919
A;Accession: T40297
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-577 <LYN>
A;Cross-references: EMBL:AL023589; PIDN:CAA19050.1; GSPDB:GN00067; SPDB:SPBC36.02c
A;Experimental source: strain 972h-; cosmid c36
C;Genetics:
A;Gene: SPDB:SPBC36.02c
A;Map position: 2
C;Superfamily: benomyl/methotrexate resistance protein
```

```
  Query Match            49.4%;  Score 41;  DB 2;  Length 577;
  Best Local Similarity  57.1%;  Pred. No. 67;
  Matches    8;  Conservative    3;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              :||::|||   ||
Db        563 FGRKIRRMSKMAVD 576


RESULT 8
A44919
GCR3 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein YM8564.07; protein YM9553.01; protein YMR125w
C;Species: Saccharomyces cerevisiae
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 29-Oct-1999
C;Accession: A44919; S53055; S54494
R;Uemura, H.; Jigami, Y.
J. Bacteriol. 174, 5526-5532, 1992
A;Title: GCR3 encodes an acidic protein that is required for expression of glycolytic ge
A;Reference number: A44919; MUID:92380925; PMID:1512188
A;Accession: A44919
A;Molecule type: DNA
A;Residues: 1-858 <UEM>
A;Cross-references: GB:D10224; NID:g464221; PIDN:BAA01076.1; PID:d1001545; PID:g464222
A;Note: sequence extracted from NCBI backbone (NCBIN:112104, NCBIP:112106)
R;Badcock, K.; Churcher, C.
submitted to the EMBL Data Library, March 1995
A;Reference number: S53055
A;Accession: S53055
A;Molecule type: DNA
A;Residues: 339-858 <BAD>
A;Cross-references: EMBL:Z48622; NID:g728663; PIDN:CAA88550.1; PID:g728664; MIPS:YMR125w
R;Lye, G.; Churcher, C.M.
submitted to the EMBL Data Library, May 1995
A;Reference number: S54014
A;Accession: S54494
A;Molecule type: DNA
A;Residues: 'MFNRKRRG',6-489 <LYE>
A;Cross-references: EMBL:Z49273; NID:g809577; PIDN:CAA89274.1; PID:g809584; MIPS:YMR125w
C;Genetics:
A;Gene: SGD:STO1; GCR3
A;Cross-references: MIPS:YMR125w; SGD:S0004732
A;Map position: 13R
C;Keywords: DNA binding; nucleus

  Query Match            49.4%;  Score 41;  DB 2;  Length 858;
  Best Local Similarity  40.0%;  Pred. No. 1e+02;
  Matches    6;  Conservative    5;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 QRYGRELRRMSDEFV 15
              :||  | | ::|:|:
Db        818 RRYSHEYRELADKFI 832


RESULT 9
F84388
hypothetical protein Vng2379h [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: F84388
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: F84388
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-84 <STO>
A;Cross-references: GB:AE004437; NID:g10581786; PIDN:AAG20474.1; GSPDB:GN00138
C;Genetics:
A;Gene: VNG2379H

  Query Match            48.2%;  Score 40;  DB 2;  Length 84;
  Best Local Similarity  66.7%;  Pred. No. 14;
  Matches    8;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          2 RYGRELRRMSDE 13
              ||| |  || ||
Db         66 RYGTEASRMRDE 77


RESULT 10
B95043
conserved hypothetical protein SP0372 [imported] - Streptococcus pneumoniae (strain 1
C;Species: Streptococcus pneumoniae
C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C;Accession: B95043
R;Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; F
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapp
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A;Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A;Reference number: A95000; MUID:21357209; PMID:11463916
A;Accession: B95043
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-109 <KUR>
A;Cross-references: GB:AE005672; PIDN:AAK74539.1; PID:g14971841; GSPDB:GN00164; TIGR
A;Experimental source: strain TIGR4
C;Genetics:
A;Gene: SP0372

  Query Match            48.2%;  Score 40;  DB 2;  Length 109;
  Best Local Similarity  45.0%;  Pred. No. 18;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    4;  Gaps    1;

Qy          1 QRYGRELRRMS----DEFVD 16
              | :|||:|  :    |||:|
Db         14 QEFGREVRGYNKVEVDEFLD 33


RESULT 11
D97913
conserved hypothetical protein spr0332 [imported] - Streptococcus pneumoniae (strain
C;Species: Streptococcus pneumoniae
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C;Accession: D97913
R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszczak, L.; Burgett, S.; DeHoff, B.S
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S
A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A;Reference number: A97872; MUID:21429245; PMID:11544234
A;Accession: D97913
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KUR>
A;Cross-references: GB:AE007317; PIDN:AAK99136.1; PID:g15457889; GSPDB:GN00174
C;Genetics:
A;Gene: spr0332

  Query Match            48.2%;  Score 40;  DB 2;  Length 113;
  Best Local Similarity  45.0%;  Pred. No. 19;
  Matches    9;  Conservative    4;  Mismatches    3;  Indels    4;  Gaps    1;

Qy          1 QRYGRELRRMS----DEFVD 16
              | :|||:|  :    |||:|
Db         18 QEFGREVRGYNKVEVDEFLD 37


RESULT 12
```

```
A75088
hypothetical protein PAB1640 - Pyrococcus abyssi (strain Orsay)
C;Species: Pyrococcus abyssi
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Aug-1999
C;Accession: A75088
R;anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A;Reference number: A75001
A;Accession: A75088
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-219 <KAW>
A;Cross-references: GB:AJ248286; GB:AL096836; NID:g5458366; PIDN:CAB50006.1; PID:e151590
A;Experimental source: strain Orsay
C;Genetics:
A;Gene: PAB1640

  Query Match             48.2%;  Score 40;  DB 2;  Length 219;
  Best Local Similarity   47.1%;  Pred. No. 37;
  Matches    8;  Conservative    4;  Mismatches    3;  Indels    2;  Gaps    1;

Qy          2 RYGRELRRMS--DEFVD 16
              ||| | :|:|  : |:|
Db        104 RYGNEFQRVSPQENFID 120


RESULT 13
E91102
probable enzyme [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C;Accession: E91102
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E91102
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <HAY>
A;Cross-references: GB:BA000007; PIDN:BAB37212.1; PID:g13363261; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3789
C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

  Query Match             48.2%;  Score 40;  DB 2;  Length 275;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              |||:|  :|  |:|
Db         35 YGRKLNALSKVFID 48


RESULT 14
A85948
probable enzyme ygfG [imported] - Escherichia coli (strain O157:H7, substrain EDL933)
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: A85948
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: A85948
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-275 <STO>
A;Cross-references: GB:AE005174; NID:g12517451; PIDN:AAG58045.1; GSPDB:GN00145; UWGP:242
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: ygfG
C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

  Query Match             48.2%;  Score 40;  DB 2;  Length 275;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              |||:|  :|  |:|
Db         35 YGRKLNALSKVFID 48


RESULT 15
F65076
hypothetical protein b2919 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C;Accession: F65076
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: F65076
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-275 <BLAT>
A;Cross-references: GB:AE000375; GB:U00096; NID:g1789282; PIDN:AAC75956.1; PID:g1789
A;Experimental source: strain K-12, substrain MG1655
C;Superfamily: naphthoate synthase; enoyl-CoA hydratase homology
F;40-192/Domain: enoyl-CoA hydratase homology <ECH>

  Query Match             48.2%;  Score 40;  DB 2;  Length 275;
  Best Local Similarity   50.0%;  Pred. No. 46;
  Matches    7;  Conservative    3;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          3 YGRELRRMSDEFVD 16
              |||:|  :|  |:|
Db         35 YGRKLNALSKVFID 48


Search completed: September 15, 2003, 17:27:03
Job time : 9.2 secs
```